Envelope Detection of Orthogonal Signals with Phase Noise [1]

Murat Azizoğlu and Pierre A. Humblet
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology

December 12, 1990

Abstract

We analyze the performance of receivers which use envelope detection at IF to detect optical signals with orthogonal modulation formats. We obtain exact closed-form expressions for the error probability conditioned on the normalized envelope. The only information necessary for obtaining the unconditional error probability is the set of tilted moments of the envelope. We then provide an approximation to this envelope which is not only accurate to the first order in phase noise strength, but also has the same range as the actual random envelope. We use this approximation to obtain the bit error performance of the three receiver models that we consider. We also provide a tight lower bound in closed-form. Finally we extend the analysis to N-ary FSK to observe the improvement due to increased bandwidth use and transmitter/receiver complexity.

[1]This research was supported by NSF under grant NSF/8802991-NCR and by DARPA under grant F19628-90-C-0002

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
12 DEC 1990		**00-12-1990 to 00-12-1990**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Envelope Detection of Orthogonal Signals with Phase Noise	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Massachusetts Institute of Technology,77 Massachusetts Avenue,Cambridge,MA,02139-4307**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**		**52**	

1 Introduction

Coherent optical communication systems provide an efficient way of sharing the huge bandwidth of single mode optical fiber among different users of the future high-speed networks. The two important reasons for this efficiency are the improved receiver sensitivity and increased channel density over direct detection systems. As a result, considerable effort has gone into developing coherent systems and into characterizing and overcoming problems to allow coherent technology to become a practical alternative to direct detection optical systems. Among major problems associated with coherent technology is the fact that the phase of the light emitted from a semiconductor laser exhibits random fluctuations due to spontaneous emissions in the laser cavity [1]. This phenomenon is commonly referred to as "phase noise". Phase noise has two important effects on the performance of optical networks. The first of these effects is the broadening of the spectrum emitted by the users of the network, which causes interchannel interference and thus necessitates wider channel spacing. This is especially dominant at low data rates where channels occupy small bandwidth. The second effect is the incomplete knowledge and the time-varying nature of the phase which makes the correct retrieval of the transmitted data bits more difficult for the receiver. While the first effect results in inefficient use of the available bandwidth, the second one causes a degradation of the bit error rate of the transmitter/receiver pairs.

It is the second effect above that we consider in this paper. This problem has received much attention in the recent years. Various performance analyses for different modulation formats and receiver structures exist in the literature [2, 3, 4, 5, 6, 7]. However, since researchers use different sets of assumptions and approximations in their work, it is difficult to reconcile their results and to agree on the quantitative effect of phase noise on coherent systems. The difficulty seems to arise from the fact that the Brownian motion model of the phase noise, which is well-tested and agreed upon, results in a random process that is both nonlinear and nonstationary. Thus, as one moves through different stages of a nontrivial receiver, the exact statistical characterization of the randomness due to phase noise becomes increasingly difficult. One needs to invoke a sequence of assumptions and approximations to overcome these difficulties. Naturally, the confidence that the real systems will operate within a reasonable margin of the predictions of the analyses depends on the number and the nature of these assumptions and approximations.

It is reasonable to expect that the modulation formats which are the least to suffer from phase noise will be those with asymmetric constellations in the signal space, e.g. on-off-keying (OOK) and frequency-shift-keying (FSK). These two modulation formats were considered recently by Foschini and his colleagues in [3]. They formulated and treated the problem with a minimal number of approximations. Their work provides a rigorous framework upon which the researchers in the field may agree and improve on. As a result, a considerable number of papers have resulted from their model in a relatively short time [7, 8, 4]. The work reported here also conforms to the receiver model of [3]. We consider orthogonal modulation formats in this paper, e.g. FSK and orthogonal polarization modulation. However the results can be generalized to obtain the performance of OOK as well. We describe the receiver models and the problem in next section. We obtain closed-form expressions for the probability of error conditional on the squared envelope of a normalized phase noisy signal, for different receiver forms. We provide an approximation which is close to the actual envelope and whose moments are readily available. We compute the probability density function of this approximate envelope from its moments using a method based on Legendre polynomials. We then use this density function to remove the conditioning on the error probability and, thus, to obtain the error performance of the systems under consideration. We also provide a lower bound to the error probability of the double-filter receiver which is very close to the actual error probability. Therefore this lower bound, which is very easy to compute, can be used to estimate the performance.

Finally, we extend the analysis to the case of N-ary FSK where we provide very tight upper and lower bounds to the bit error probability.

2 Problem Description

The received optical signal is first processed by an optical heterodyne receiver shown in Figure 1 for FSK signals. Polarization modulation would require a polarization beam splitter and a pair of photodetectors. Optical heterodyning transforms the signal from optical frequency to intermediate frequency. The IF signal output $r(t)$ is to be further processed by the IF receiver. This IF processing is the main focus of this paper.

The IF receiver structure that we consider in this paper is shown in Figure 2 where the incoming signal is corrupted by phase noise and additive noise. The front end of this receiver is the standard quadrature receiver that performs envelope detection

of FSK signals. The correlators and integrators perform a bandpass filtering to limit the noise power corrupting the signal. Since the integrators integrate their inputs over a duration T' the effective bandwidth of the filter is $1/T'$. For a uniformly distributed phase uncertainty which is constant over the bit duration the optimum value of T' is the bit duration T. However, when the signal is corrupted by phase noise, the spectrum of the signal is broadened. Therefore a wider filter bandwidth, equivalently smaller integration times, may be necessary. For analytical convenience, we only consider the case where the ratio T/T' is a positive integer M as in [3].

The outputs of the in-phase and quadrature branch integrators in Figure 2 are squared and then added to perform the envelope detection. The remaining processing depends on the value of M. For $M = 1$, the adder outputs are sampled at the end of the bit duration and the two values are compared to reach a decision about the transmitted data bit. We call the receiver with $M = 1$ a *single sample receiver* since only one sample is taken during a bit duration. For $M \geq 2$, the adder outputs are sampled every T' seconds resulting in M samples per bit. The way these samples are processed depends on the complexity that one desires. One simple strategy is to discard all but the last one of these samples and to use the last sample as in the single sample receiver. A more efficient way is to average these M samples and compare the two averages. This averaging is not the optimal processing of the samples. However since it can be performed by a lowpass filter it is a practical reception strategy to implement.

We classify the reception strategies above into three categories. The first is the single sample receiver, the second one is the multisample receiver with single filtering and the third one is the multisample receiver with double filtering. These receiver structures were first suggested by Kazovsky et. al. in [10] for ASK multiport homodyne receivers. In the framework of [10], the single sample receiver is the conventional matched-filter receiver, the single-filter receiver is the conventional receiver with a widened filter, and the double filter receiver is the wide-band filter-rectifier-narrowband filter (WIRNA) structure. It was also noticed in [10] that if the M samples of the double filter receiver are viewed as coming from M distinct channels (or from a fast frequency-hopped spread spectrum system), then the receiver is functionally equivalent to the noncoherent receiver for multichannel signaling treated in [11]. Some of our results, in particular the conditional error probability given by Equation (19), may be obtained by using the results of [11]. However, we will proceed independently for the sake of completeness.

2.1 Single Sample Receiver

We first consider the single sample receiver. This receiver is the optimal receiver when the phase of the received signal has uniform distribution and is time-invariant. The error probability for this case has a simple closed-form ($1/2\ e^{-E_b/2N_0}$) [12]. We will see that a similar (but more complicated) result can be obtained in the phase noisy case.

We assume that the received IF signal is the FSK signal corrupted by additive noise and phase noise:

$$r(t) = A\cos(2\pi f_i t + \theta(t)) + n(t) \quad i = 0, 1 \tag{1}$$

where f_0 and f_1 are the frequencies for "0" and "1" respectively. The phase noise $\theta(t)$ is a Brownian motion process described by

$$\theta(t) = 2\pi \int_0^t \mu(\tau)\, d\tau \tag{2}$$

with $\mu(\tau)$ being a zero-mean white Gaussian process with two-sided spectral density $\beta/2\pi$. Then $\theta(t)$ is a zero-mean Gaussian process with zero-mean and variance $2\pi\beta t$; β is the combined linewidth of the source and local oscillator lasers. The additive noise $n(t)$ in (1) is a white Gaussian process with two-sided spectral density $N_0/2$. We assume that the difference between f_1 and f_0 is much larger than the filter bandwidth $1/T'$, so that when a "0" is transmitted, the integrator outputs of the lower branch integrators in Figure 2a contain no signal component, and vice versa. With this assumption, when a "0" is transmitted the sampler outputs at time T can be written as

$$\begin{aligned} Y_0 &= \left| \frac{A}{2} \int_0^T e^{j\theta(t)}\, dt + n_{c0} + j n_{s0} \right|^2 \\ Y_1 &= |n_{c1} + j n_{s1}|^2 \end{aligned} \tag{3}$$

where n_{ci} and n_{si} are statistically independent and identically distributed (i.i.d.) Gaussian random variables with zero-mean and variance $\sigma^2 = N_0 T/4$. Due to symmetry, the error probability is given by

$$P_e = \Pr(Y_0 \leq Y_1)\ . \tag{4}$$

We will now condition the error probability on the amplitude of the phase noise integral in (3). We define

$$Y \triangleq \left| \frac{A}{2} \int_0^T e^{j\theta(t)}\, dt \right| . \tag{5}$$

Then the conditional density of Y_0 can be easily found as[2]

$$p(y_0|Y) = \frac{1}{2\sigma^2} e^{-(y_0+Y^2)/2\sigma^2} I_0\left(\frac{Y\sqrt{y_0}}{\sigma^2}\right) \tag{6}$$

while Y_1 has an exponential density given by

$$p(y_1) = \frac{1}{2\sigma^2} e^{-y_1/2\sigma^2} . \tag{7}$$

Therefore the conditional error probability comes as a result of the classical problem of incoherent detection of orthogonal signals as [12]

$$P_e(Y) = \frac{1}{2} e^{-Y^2/2\sigma^2} . \tag{8}$$

Let's now express Y in a different form. From (5), we have

$$Y = \frac{AT}{2} \left| \int_0^1 e^{j\theta(Tu)}\, du \right| .$$

Since $\theta(Tu)$ has variance $2\pi\beta Tu$, Y can be written in terms of standard Brownian motion $\psi(t)$ (i.e. that with variance t) as

$$Y = \frac{AT}{2} \left| \int_0^1 e^{j\sqrt{\gamma}\psi(t)}\, dt \right|$$

where $\gamma \triangleq 2\pi\beta T$. Now if we define the random variable $X(\gamma)$ as

$$X(\gamma) \triangleq \left| \int_0^1 e^{j\sqrt{\gamma}\psi(t)}\, dt \right|^2 \tag{9}$$

(8) can be written in terms of X as

$$P_e(X) = \frac{1}{2} e^{-\xi X(\gamma)/2} \tag{10}$$

where $\xi = A^2T/2N_0$ is the signal-to-noise ratio. Therefore P_e is given by

$$P_e = E_X\left[\frac{1}{2} e^{-\xi X(\gamma)/2}\right] . \tag{11}$$

In the case with no phase noise ($\gamma = 0$), this reduces to $1/2e^{-\xi/2}$ as it should.

An immediate implication of (11) is that there does not exist a bit-error-rate-floor, i.e. a nonzero error probability that exists even when ξ tends to ∞, for the

[2]For notational convenience we indirectly address the random variable that a density function refers to with its argument, i.e., $p(y)$ stands for $p_Y(y)$ in the more standard notation.

single sample receiver, provided that the probability density function of $X(\gamma)$ does not contain a delta function at the origin [3].

To determine P_e from (11) we have to obtain either the probability density function or the moment generating function of X. Therefore the random variable X plays a key role in the performance analysis of the single sample receiver. In the next section, we will show that this is also true for the multisample receivers that we consider.

2.2 Multisample Receivers

In the previous section, we mentioned that the receiver may use a shorter integration time T' than the full bit duration T. The advantage of this scheme is that the phase is not allowed to wander too much so that the amount of collected signal energy may not drop below undesirable levels. Equivalently, the bandwidth of the IF filter is increased to pass more of the signal power. If only the last of the M samples, the one at $t = T$, is used, then the received signal is effectively processed only during the interval $(T - T', T)$, hence there is a drop in the received signal-to-noise ratio by a factor of $1/M$. However, since the phase is allowed to change in an interval of length T', the effective value of γ also drops by a factor of $1/M$. (This can be observed by repeating the calculations of the previous section when T is replaced by T'.) Therefore the probability of error for the multisample receiver with single filter is given by

$$P_e = E_X \left[\frac{1}{2} e^{-\xi X(\gamma/M)/2M} \right] . \tag{12}$$

An inspection of (12) shows that when the signal-to-noise ratio and/or the phase noise strength is small, the optimal value of M will be small. However, for large signal-to-noise ratios and large phase noise values, it may be worthwhile to sacrifice from the signal energy while alleviating the effect of phase noise.

The single-filter receiver uses only one of the M available samples. Thus the single-filter receiver is, in fact, a single sample receiver with widened front end filter bandwidth. The efficiency of this receiver would be improved if all the M samples were used in the decision process. A simple processing of these samples is averaging, which may be accomplished by low-pass-filtering the sum of the outputs of the square-law devices [3]. The result is the multisample double-filter receiver that we analyze below. Note that this receiver is not the (still unknown) receiver which minimizes the

[3] Note that the inherent error floor in FSK due to the finite frequency separation [13, 5] is neglected in our model.

probability of error. It is closely related to the optimal receiver for the case where the phase noise process $\theta(t)$ is piecewise constant and takes M independent uniformly distributed values during a bit interval. In this case, the optimal receiver averages the square root of the samples when the signal-to-noise ratio is high.

In the case of a "0" being transmitted, the lowpass filter outputs will be

$$Y_0 = \sum_{k=1}^{M} |x(k) + n_{c0}(k) + jn_{s0}(k)|^2 \tag{13}$$

$$Y_1 = \sum_{k=1}^{M} |n_{c1}(k) + jn_{s1}(k)|^2 \tag{14}$$

where $x(k)$ is given by

$$x(k) = \frac{A}{2} \int_{(k-1)T'}^{kT'} e^{j\theta(t)}\, dt$$

and all the additive noise samples are i.i.d. zero-mean Gaussian with variance

$$\alpha^2 = N_0 T'/4 \; .$$

Since $\theta(t)$ has independent increments, $x(k)$ are statistically independent. Furthermore, since $\theta(t)$ has stationary increments, $|x(k)|$ will be identically distributed[4]. Then each of the terms in (13) and (14) obey the distributions in (6) and (7) respectively. (Y in (6) is replaced by $|x(k)|$.) In particular, since each term in (14) has i.i.d. exponential distribution, Y_1 has a Gamma distribution with parameters M and $1/2\alpha^2$. It is convenient to normalize Y_0 and Y_1 via

$$V_i = \frac{Y_i}{2\alpha^2} \quad i = 0, 1 \; .$$

Then V_1 has the density function

$$p(v_1) = \frac{v_1^{M-1}}{(M-1)!} e^{-v_1} \quad v_1 \geq 0 \; . \tag{15}$$

Due to symmetry, the probability of error is given by

$$\begin{aligned} P_e &= \Pr(V_0 \leq V_1) \\ &= \int_0^{\infty} \sum_{k=0}^{M-1} \frac{v_0^k}{k!} e^{-v_0} p(v_0)\, dv_0 \end{aligned} \tag{16}$$

[4]This also implies that the envelopes preserve their statistical properties from one bit to another. So we can consider the first bit in the performance analysis without loss of generality.

where $p(v_0)$ is the density function of V_0. This density appears to be difficult to obtain in closed-form. Therefore, we will first condition (16) on the phase noisy envelopes. Given $\{x(k) : 1 \leq k \leq M\}$, V_0 is the sum of squares of $2M$ independent Gaussian random variables, the density of which is given by [12]

$$p\left(v_0|\{x(k)\}\right) = \left(\frac{v_0}{r}\right)^{(M-1)/2} e^{-(v_0+r)} I_{M-1}\left(\sqrt{4rv_0}\right) \quad v_0 \geq 0 \tag{17}$$

where $I_{M-1}(\cdot)$ is the modified Bessel function of order $M-1$ and r is defined as

$$r = \frac{1}{2\alpha^2}\sum_{k=1}^{M} |x(k)|^2 \ .$$

Note that the conditional density of V_0 depends on $\{x(k)\}$ only through the sum of their squared amplitudes. This enables us to proceed further and to express the error probability conditioned on r as

$$\begin{aligned} P_e(r) &= \Pr(V_0 \leq V_1|r) \\ &= \sum_{k=0}^{M-1} \frac{1}{k!} \int_0^\infty e^{-x} x^k \left(\frac{x}{r}\right)^{(M-1)/2} e^{-(x+r)} I_{M-1}\left(\sqrt{4rx}\right)\, dx \ . \end{aligned} \tag{18}$$

It is shown in Appendix A that $P_e(r)$ can be expressed as

$$P_e(r) = \frac{e^{-r/2}}{2^M} \sum_{n=0}^{M-1} \frac{(r/2)^n}{n!} \sum_{k=n}^{M-1} \binom{k+M-1}{k-n} \frac{1}{2^k} \ . \tag{19}$$

Now we relate r to the envelope of the standard Brownian motion using the techniques of the previous section. From the definition of r and $x(k)$ we have

$$\begin{aligned} r &= \frac{A^2}{8\alpha^2} \sum_{k=1}^{M} \left| \int_{(k-1)T/M}^{kT/M} e^{j\theta(t)} dt \right|^2 \\ &= \frac{A^2}{8\alpha^2} \left(\frac{T}{M}\right)^2 \sum_{k=1}^{M} \left| \int_{k-1}^{k} e^{j\sqrt{\gamma/M}\psi(t)} dt \right|^2 \end{aligned}$$

where $\psi(t)$ is the standard Brownian motion. Defining

$$X_k(\gamma) = \left| \int_{k-1}^{k} e^{j\sqrt{\gamma}\psi(t)} dt \right|^2$$

and using the definitions of the additive noise variance α^2 and the signal-to-noise ratio (SNR) ξ, we obtain

$$r = \xi \frac{1}{M} \sum_{k=1}^{M} X_k(\gamma/M) \ . \tag{20}$$

Therefore r is the SNR weighted by the sample average of the M normalized phase noisy envelope squares, it can be viewed as a phase noisy signal-to-noise ratio. As $M \rightarrow \infty$, the sample average tends to the mean of X_k with probability 1, by the law of large numbers. Also, the mean of $X_k(\gamma/M)$ tends to 1 as $M \rightarrow \infty$. Therefore r tends to ξ, which is the maximum value it can assume since $X_k \leq 1$ by Schwartz inequality.

The conditional error probability given by (19) is shown in Figure 3 as a function of r. As expected $P_e(r)$ decreases as r gets larger while M is fixed. However, for a fixed value of r, $P_e(r)$ increases with M. Then, there is a tradeoff between the collected energy of the phase noisy signal and the additive noise energy corrupting the signal as M changes. There must exist an optimum value of M which balances the two conflicting noise effects. This behavior is also observed in [3].

We can readily obtain a lower bound to the error probability without calculating the density of $X(\gamma)$. We show in Appendix B that $P_e(r)$ as given by (19) is a convex $\cup$ function of r for any value of M. The unconditional error probability is the expected value of $P_e(r)$; thus replacing r by its mean results in a lower bound by Jensen's inequality [14]. Using (20), one has

$$P_e \geq P_e(E[X(\gamma/M)]) \ . \tag{21}$$

The expected value of $X(\gamma)$ will be found in the next section. Figure 4 shows the lower bound in (21) with the use of (34) of the next section. At each value of SNR ξ and the phase noise strength γ, the lower bound has been minimized with respect to M.

The overall error probability P_e is obtained from $P_e(r)$ by removing the conditioning over r. If $p_a(r)$ denotes the probability density function of r, then we have

$$P_e = \int P_e(r) p_a(r)\, dr \ . \tag{22}$$

Since r is a weighted average of i.i.d. random variables X_k as given by (20), the density of r is given in terms of the density of X_k by an M-fold convolution. Then, we see that the density of the phase noisy squared-envelope is again the only information we need to obtain the error probability. The convolutions must be performed for each M since the individual densities to be convolved correspond to the envelope with phase noise strength γ/M. Therefore the computation becomes very costly with many convolutions to be performed. We describe an alternative method which eliminates

the need for convolutions. From (19), the error probability can be written as

$$P_e = E_r \left[\sum_{n=0}^{M-1} a_n \left(\frac{r}{2} \right)^n e^{-r/2} \right] \tag{23}$$

where the coefficients a_n are defined as

$$a_n = \frac{1}{2^M n!} \sum_{k=n}^{M-1} \binom{k+M-1}{k-n} \frac{1}{2^k} . \tag{24}$$

Using the sample average expression for r given in (20) one gets

$$P_e = \sum_{n=0}^{M-1} a_n \left(\frac{\xi}{2M} \right)^n E \left[\left(\sum_{k=1}^{M} X_k(\gamma/M) \right)^n \exp\left(-\frac{\xi}{2M} \sum_{k=1}^{M} X_k(\gamma/M) \right) \right] .$$

Finally expanding the expression above in a multinomial sum and using the independence of the X_k we obtain

$$P_e = \sum_{n=0}^{M-1} a_n \left(\frac{\xi}{2M} \right)^n \sum_{\substack{k_1,\ldots,k_M \\ k_1+\cdots+k_M=n}} \frac{n!}{k_1! \cdots k_M!} \prod_{i=1}^{M} \alpha(k_i) \tag{25}$$

where

$$\alpha(k) = E \left[X^k(\gamma/M) \exp\left(-\frac{\xi}{2M} X(\gamma/M) \right) \right] . \tag{26}$$

The need for convolutions is eliminated by this expansion. All that is necessary about the envelope is a set of tilted moments $\alpha(k)$. This reduces the computations considerably as will be seen in Section 5. Further computational savings can be obtained by observing that the inner summation in (25) contains many identical terms: any permutation of an M-tuple $(k_1, k_2, \ldots, k_M)$ results in the same term. Therefore the summation may be limited to ordered tuples with $k_1 \geq k_2 \geq \cdots \geq k_M$ with the introduction of a scaling factor that counts the number of permutations. This scaling factor depends on how distinct the entries of the tuple are. If all the entries in a tuple $\vec{k}$ are distinct, then $\vec{k}$ has $M!$ permutations. If an entry k_i is repeated $r(k_i)$ times in $\vec{k}$, then every distinct permutation has $r(k_i)!$ copies. Then, the scaling factor with ordered indices is

$$N(\vec{k}) = \frac{M!}{\prod_i r(k_i)!}$$

where the product is taken over i for which k_i are distinct.

3 Phase Noisy Envelope and An Approximation

In the previous sections, we have seen that the squared-envelope of a phase noisy sinusoidal plays a critical role in the performance analysis of the incoherent IF reception of heterodyne FSK system. This squared-envelope is the random variable X defined in (9) which we repeat here for convenience:

$$X = \left| \int_0^1 e^{j\sqrt{\gamma}\psi(t)}\, dt \right|^2 \tag{27}$$

where we have removed the explicit argument γ for notational simplicity. Due to its importance in the performance analysis, the determination of the probability density of X, or equivalently its moment generating function, has received considerable attention. Foschini and his coworkers provided the first comprehensive treatment of this problem in [3] and [9]. We summarize their approach below, since it is directly related to our subsequent analysis.

Foschini et. al. expand the complex exponential into its power series and retain the first order γ terms. The resulting approximation, X_L, is linear in γ and is given by

$$X_L = 1 - \gamma \left[\int_0^1 \psi^2\, dt - \left(\int_0^1 \psi\, dt \right)^2 \right] . \tag{28}$$

(We use the subscript L to emphasize the linear nature of this approximation.) Foschini et. al. expand $\psi(t)$ in Fourier cosine series on $(0,1)$ and obtain

$$X_L = 1 - \gamma \mathbf{z}^T D \mathbf{z} \tag{29}$$

where $\mathbf{z}$ is an infinite dimensional vector of independent, identically distributed Gaussian random variables with zero-mean and unit variance, and D is an infinite dimensional diagonal matrix with $d_{ii} = 1/(i\pi)^2$. Thus, $\mathbf{z}^T D \mathbf{z}$ is a compact notation for $\sum_{i=1}^{\infty} z_i^2/(i\pi)^2$. The moment generating function of X_L is then found as [9]

$$M_{X_L}(s) \triangleq E\left(e^{sX_L}\right) = e^s \sqrt{\frac{\sqrt{2\gamma s}}{\sinh \sqrt{2\gamma s}}} \tag{30}$$

and is inverse Laplace-transformed to obtain the density function of X in semi-closed form.

Equation (30) can be used in conjunction with (11) to obtain the error probability for the single sample receiver with the linear approximation. The result is obtained as

$$P_e = \begin{cases} \frac{1}{2} e^{-\xi/2} \left(\sqrt{\xi\gamma} / \sin \sqrt{\xi\gamma} \right)^{1/2} & \xi\gamma < \pi^2 \\ \infty & \xi\gamma \geq \pi^2 \end{cases} \tag{31}$$

where the first case follows via $\sinh(jx) = j\sin(x)$, while the second case follows from the derivation of (30) in [9]. It is seen that this approximation fails to give sensible results at high SNR values. This is due to the fact that although X can take values only between 0 and 1, X_L can become negative. In fact, the density of X_L has a nonzero tail for negative arguments, which becomes increasingly dominant as γ increases. According to (31), this negative tail of the density function $p_{X_L}(x)$ decays exponentially, as $\exp(\pi^2 x/2\gamma)$ for $x < 0$, so that when $\xi > \pi^2/\gamma$ the integral of $e^{-\xi x/2} p_{X_L}(x)$ diverges. This problem is not observed in the numerical results of [3] due to the truncation of the negative tail of the probability density function. However, the truncation weakens the analytical basis of the approximation: the truncated density function is no longer a first order approximation to that of the actual envelope, this introduces an uncertainty on the error probability results.

We seek a better approximation to the random variable X by forcing the range of the approximate random variable, say $\bar{X}$, to match the range of X, namely $(0,1)$. Therefore we require a good approximation to satisfy

1. $O \leq \bar{X}(\gamma) \leq 1$ for all $\gamma \geq 0$ and all Brownian sample paths $\psi(t)$,
2. $\bar{X}(\gamma)$ and $X(\gamma)$ match to the first order in γ for all $\psi(t)$.

The range constraint above suggests that $\bar{X}(\gamma)$ can be written as

$$\bar{X}(\gamma) = \exp\left[-G(\sqrt{\gamma}\psi)\right]$$

where G is a functional whose domain is the set of functions defined on $[0,1]$ which vanish at 0, and whose range is $[0,\infty)$. We force G to satisfy

1. $G(f) \geq 0$ for all $f(t)$, with equality if and only if $f(t) \equiv 0$
2. $G(f) = G(-f)$ for all $f(t)$,

where the second property is due to the fact that the two sample paths $\psi(t)$ and $-\psi(t)$ result in the same value of $X(\gamma)$. For $X(\gamma)$ and $\bar{X}(\gamma)$ to be identical to the first order in γ, $X(0) = \bar{X}(0)$ and $\bar{X}'(0) = X'(0)$ must be satisfied. where 'prime' denotes the first derivative with respect to γ. We state one functional $G(f)$ that satisfies these requirements in the following lemma.

Lemma 1: If $G(f) = \int_0^1 f^2\,dt - \left(\int_0^1 f\,dt\right)^2$, then $X(\gamma)$ and $\bar{X}(\gamma)$ are identical to the first order in γ.

Proof : Since $X(0) = 1$, $G(0) = 0$ is needed for $X(0) = \bar{X}(0)$. On the other hand, it is straightforward, but somewhat lengthy, to show that

$$X'(0) = \left(\int_0^1 \psi \, dt\right)^2 - \int_0^1 \psi^2 \, dt \; .$$

For a G that satisfies $G(0) = 0$,

$$\bar{X}'(0) = -\frac{\partial G(\sqrt{\gamma}\psi)}{\partial \gamma}|_{\gamma=0} \; .$$

It is easily seen that the functional given in the statement of the lemma satisfies both of the requirements. □

Note that the functional of the Lemma 1 satisfies the positivity and evenness conditions as well. Therefore the random variable X_E defined as

$$X_E(\gamma) \triangleq \exp(-\gamma \mathbf{z}^T D \mathbf{z}) \tag{32}$$

with $\mathbf{z}$ and D as defined previously, promises to be a good approximation to X. We use the subscript E to denote the exponential nature of this approximation.

To summarize, we have obtained an approximation that retains the desirable feature of that of Foschini et. al. while having the additional feature that it takes values that are in the same range as the original random variable.

A nice feature of the random variable X_E is that its moments are easily obtained from (30) as

$$\begin{aligned} \mu(t) &\triangleq E(X_E^t) \\ &= \sqrt{\frac{\sqrt{2\gamma t}}{\sinh(\sqrt{2\gamma t})}} \end{aligned} \tag{33}$$

for all real $t \geq -\pi^2/2\gamma$. For $t \leq -\pi^2/2\gamma$, the t'th moment of X_E is infinite.

An indication of how well the two approximations model X is the behavior of the first two moments as γ varies. The moments of X_E are already found in (33). We now evaluate the first two moments of X. Rewriting (27) as a double integral and taking the expectation we obtain

$$E(X) = \int_0^1 \int_0^1 E\left\{\exp\left[j\sqrt{\gamma}\left(\psi(t) - \psi(s)\right)\right]\right\} \, dt \, ds \; .$$

Now we use the fact that $\sqrt{\gamma}\left(\psi(t) - \psi(s)\right)$ is Gaussian with zero-mean and variance $\gamma \left|t - s\right|$ and get

$$\begin{aligned} E(X) &= \int_0^1 \int_0^1 \exp\left(\gamma \left|t - s\right| / 2\right) \, dt \, ds \\ &= \frac{4}{\gamma}\left[1 - \frac{2}{\gamma}\left(1 - e^{-\gamma/2}\right)\right] \; . \end{aligned} \tag{34}$$

The calculation of $E(X^2)$ follows along the same lines: we first express X^2 as a four-fold integral, take the expectation and express the integrand as a Gaussian characteristic function while observing the dependencies of the random variables. We finally obtain

$$E\left(X^2\right) = \frac{16}{\gamma^2}\left[2 - \frac{15}{\gamma} + \frac{87}{2\gamma^2} - \left(\frac{20}{3\gamma} + \frac{392}{9\gamma^2}\right)e^{-\gamma/2} + \frac{1}{18\gamma^2}e^{-2\gamma}\right] . \tag{35}$$

The first two moments of X as given by Equations (34) and (35) are in agreement with previous calculations of the mean and the variance in [10] and [13].

Finally we calculate the first two moments of X_L (without truncation). The first moment is obtained as

$$\begin{aligned} E(X_L) &= 1 - \gamma E\left(\mathbf{z}^T D\mathbf{z}\right) \\ &= 1 - \gamma \sum_{n=1}^{\infty} \frac{1}{(n\pi)^2} \\ &= 1 - \frac{\gamma}{6} \end{aligned} \tag{36}$$

and the second moment is obtained as

$$\begin{aligned} E\left({X_L}^2\right) &= 1 - 2\gamma E\left(\mathbf{z}^T D\mathbf{z}\right) + \gamma^2 E\left((\mathbf{z}^T D\mathbf{z})^2\right) \\ &= 1 - 2\gamma \sum_{n=1}^{\infty} \frac{1}{(n\pi)^2} + \gamma^2 \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{E(z_n^2 z_m^2)}{(n\pi)^2(m\pi)^2} \\ &= 1 - \frac{\gamma}{3} + \gamma^2 \sum_{n=1}^{\infty} \frac{3}{(n\pi)^4} + \gamma^2 \sum_{n\neq m}\sum_{m=1}^{\infty} \frac{1}{(n\pi)^2}\frac{1}{(m\pi)^2} \\ &= 1 - \frac{\gamma}{3} + \frac{\gamma^2}{20} \end{aligned} \tag{37}$$

where we have used $E(z_n^4) = 3$ and $\sum_{n=1}^{\infty} n^{-4} = \pi^4/90$ [15]. The first two moments of X, X_E and X_L as given by Equations (33)-(37) are shown in Figure 5 as functions of γ. It can be seen from the figure that the moments of X_L agree with those of the actual random variable X only for very small γ while the moments of X_E are very close to the actual moments for all γ. Therefore we expect the exponential approximation to be an accurate one.

4 Density of X_E

The moments of X_E as given by (33) provides complete information about its statistical behavior. In principle, one could compute the moment generating function of

X_E from the moments $\{\mu(n)\}$ as

$$F(s) \triangleq E\left(e^{-sX_E}\right) = \sum_{n=0}^{\infty} \frac{(-s)^n}{n!}\mu(n)$$

and find the density function as the inverse-Laplace transform. However, since the series above is slowly converging, especially for large s, this is not an computationally attractive approach. Instead, we use an indirect approach which is based on orthogonal polynomials for finding the density. We first prove the following about the value of the density function at the origin.

Lemma 2: Let $q(x)$ denote the probability density function of X_E. Then $q(0)$ has the following property:

$$q(0) = \begin{cases} 0 & \text{if } \gamma < \pi^2/2 \\ \infty & \text{if } \gamma > \pi^2/2\ . \end{cases} \tag{38}$$

Proof : We prove the first part of the lemma by using Chernoff bound [16]. We have

$$q(0) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \Pr\left(0 \le X_E < \epsilon\right)\ .$$

From the definition of X_E the probability above is

$$Pr\left(0 \le X_E < \epsilon\right) = \Pr\left(\mathbf{z}^T D\mathbf{z} > -\frac{1}{\gamma}\ln\epsilon\right)\ .$$

We now bound the right hand side by using the Chernoff bound on $\mathbf{z}^T D\mathbf{z}$ via (33) to obtain

$$Pr\left(0 \le X_E < \epsilon\right) \le \epsilon^{s/\gamma}\left(\frac{\sqrt{2s}}{\sin(\sqrt{2s})}\right)^{1/2} \quad \text{for } 0 \le s \le \pi^2/2\ .$$

Dividing both sides by ϵ and taking the limit we obtain

$$q(0) \le \lim_{\epsilon \to 0} \epsilon^{s/\gamma - 1}\left(\frac{\sqrt{2s}}{\sin(\sqrt{2s})}\right)^{1/2} \quad \text{for } 0 \le s \le \pi^2/2\ .$$

The limit above is 0 if $s/\gamma > 1$, which is possible for some s in the allowed range if $\gamma < \pi^2/2$. Since $q(x) \ge 0$, $q(0) = 0$.

For the second part of the proof, we use the fact that each of the random variables in the series of $\mathbf{z}^T D\mathbf{z}$ are nonnegative to obtain the following lower bound.

$$Pr\left(0 \le X_E < \epsilon\right) = \Pr\left(\mathbf{z}^T D\mathbf{z} > -\frac{1}{\gamma}\ln\epsilon\right) \ge \Pr\left(\frac{1}{\pi^2}z_1^2 > -\frac{1}{\gamma}\ln\epsilon\right)\ .$$

The rightmost probability is $2Q(\sqrt{-\pi^2 \ln \epsilon/\gamma})$ where $Q(x)$ is the complementary distribution function for the unit normal random variable. Proceeding as before we obtain

$$q(0) \geq \lim_{\epsilon \to 0} \frac{2}{\epsilon} Q\left(\sqrt{-\frac{\pi^2}{\gamma} \ln \epsilon}\right) .$$

Finally we use $Q(x) \sim \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}$ as $x \to \infty$ [17] to obtain

$$q(0) \geq \sqrt{\frac{2\gamma}{\pi^3}} \lim_{\epsilon \to 0} \frac{\epsilon^{\pi^2/2\gamma - 1}}{\sqrt{-\ln \epsilon}} .$$

The limit above is infinite if $\gamma > \pi^2/2\gamma$, this completes the proof. □

The lemma above indicates that the density of X_E changes its form as γ exceeds a critical value, $\pi^2/2$. The left tail of the density, i.e. $q(x)$ for small x, strongly affects the probability of error, since P_e is the integral of the product of $q(x)$ (or convolution of $q(x)$ with itself) and some rapidly decaying function such as $e^{-\xi x/2}$. Therefore the behavior of $q(x)$ for $x \ll 1$ is of special interest for accurate calculation of error probability. The following lemma indicates that this behavior is polynomial.

Lemma 3: $q(x) = O\left(x^{\frac{\pi^2}{2\gamma}-1}\right)$ for x small, where $f(x) = O(g(x))$ is equivalent to $\lim_{x \to 0} f(x)/g(x) = c$ for a positive finite c.

Proof : From (33) we have

$$\int_0^1 x^t q(x)\, dx = \begin{cases} \left(\sqrt{2\gamma t}/\sinh(\sqrt{2\gamma t})\right)^{1/2} & t > 0 \\ \left(\sqrt{2\gamma|t|}/\sin(\sqrt{2\gamma|t|})\right)^{1/2} & 0 > t > -\pi^2/2\gamma \\ \infty & t < -\pi^2/2\gamma . \end{cases}$$

Since $\int_0^\delta x^{-\beta}\, dx$ converges for $\beta < 1$ and diverges for $\beta \geq 1$ for any $\delta > 0$, we must have

$$x^{-\frac{\pi^2}{2\gamma}} q(x) = c(x)\frac{1}{x}$$

for some function $c(x)$ with finite and positive $c(0)$, where we have assumed that $q(x)$ is continuous at $x = 0$. Then the claim is proved. □

This lemma is in agreement with the previous lemma in predicting the value of $q(0)$. Moreover, it suggests that for accurate calculation of $q(x)$ for small x it is better to calculate the tilted density $q(x)/x^{\frac{\pi^2}{2\gamma}-1}$. Before we describe a method to perform this calculation, we exploit a useful property of $X_E(\gamma)$ that will enable us to compute the density function for a single value of γ, say $\gamma = 1$, and to obtain all other density functions from this computed one. From the definition of $X_E(\gamma)$ we have

$$X_E(\gamma)^{1/\gamma} = e^{-\mathbf{z}^T D \mathbf{z}} .$$

Then for any positive γ_1 and γ_2, the associated X_E's satisfy

$$X_E(\gamma_1)^{1/\gamma_1} = X_E(\gamma_2)^{1/\gamma_2}$$

which specifies a one-to-one mapping from $X_E(\gamma_1)$ to $X_E(\gamma_2)$ via

$$g(x) = x^{\gamma_2/\gamma_1} \ .$$

Then the densities of the two random variables are related via [18]

$$q_{\gamma_2}(x) = \frac{q_{\gamma_1}\left(g^{-1}(x)\right)}{|g'\left(g^{-1}(x)\right)|}$$

where we introduced a subscript to the density function to emphasize its dependence on γ. Upon manipulation, we obtain

$$q_{\gamma_2}(x) = \alpha x^{\alpha-1} q_{\gamma_1}(x^\alpha) \tag{39}$$

where $\alpha = \gamma_1/\gamma_2$.

The calculation of $q(x)$ for a given γ, as will be seen shortly, is a computationally intensive effort. Therefore the simple relation (39) introduces a big saving in the required numerical work.

The moments of a random variable are the projections of its density function on a sequence of polynomials. Therefore, the density can be reconstructed from these moments via a complete orthonormal polynomial basis. The even (or odd) indexed Legendre polynomials constitute such a basis on the interval $(0,1)$ [19]. Any piecewise continuous function $f(x)$ on $(0,1)$ that has left and right derivatives at every point in the interval can be expressed as

$$f(x) = \sum_{\substack{n=0 \\ n \text{ even}}}^{\infty} (2n+1) F_n P_n(x) \tag{40}$$

where $P_n(x)$ is the n'th Legendre polynomial and F_n is given by

$$F_n = \int_0^1 f(x) P_n(x)\, dx \ . \tag{41}$$

Specializing to the case of

$$f(x) = \frac{q(x)}{x^h}$$

where we have defined

$$h \triangleq \frac{\pi^2}{2\gamma} - 1$$

we obtain

$$F_n = E\left[X_E^{-h} P_n(X_E)\right] . \tag{42}$$

To obtain the coefficients F_n in terms of the moments, we first expand $P_n(x)$ in its Taylor series as [19]

$$P_n(x) = \sum_{\substack{k=0 \\ k \text{ even}}}^{n} a_{nk} x^k \tag{43}$$

where the coefficients a_{nk} are given by

$$a_{nk} = \frac{1}{2^n}(-1)^{(n-k)/2} \binom{n}{(n-k)/2} \binom{n+k}{k} . \tag{44}$$

Now we can compute the expectation in (42) in terms of the moments given by (33) as

$$F_n = \sum_{\substack{k=0 \\ k \text{ even}}}^{n} a_{nk} \mu(k-h) \tag{45}$$

and, consequently, the density function as

$$q(x) = x^h \sum_{\substack{n=0 \\ n \text{ even}}}^{\infty} (2n+1) P_n(x) \sum_{\substack{k=0 \\ k \text{ even}}}^{n} a_{nk} \mu(k-h) . \tag{46}$$

The convergence of the series in (46) is quite fast. Note that the moments always remain finite since $k - h > -\pi^2/2\gamma$ for all $k \geq 0$. It is also worth mentioning that although the formula above remains valid for any value of h, in particular for $h = 0$ which corresponds to the direct computation of $q(x)$ without tilting, the value that we use is the most appropriate one because it captures the polynomial behavior of $q(x)$ in the small x region.

The results of the computation is shown in Figure 6 for various γ values. As expected the density is more concentrated around unity for small γ. As γ increases, the density gets flat until $\gamma = \pi^2/2$ and becomes peaked at $x = 0$ thereafter. The density function is observed to have a sharp decay around its $x = 1$ tail. In Appendix C, we show that the nature of this decay is exponential in $\gamma/(1-x)$, provided that $q(x)$ is bounded and analytic. (Boundedness holds for $\gamma < \pi^2/2$ and analyticity is assumed to hold.)

The density function of X_E can also be calculated with another method. The distribution function of the random variable $\mathbf{z}^T D \mathbf{z}$ is given in [3], from which the distribution function of X_E can be directly computed, the density function is then

obtained by differentiation. It was observed that this method yields results which are in perfect agreement with Figure 6.

In the next section, we will use the density we have obtained to evaluate the performance of the receivers under consideration.

5 Results and Discussion

The error probability for the single sample receiver is given by (11) as

$$P_e = \frac{1}{2}\int_0^1 e^{-\xi x/2} q_\gamma(x)\, dx\ .$$

We have computed the error probability using this integral. The result of this computation is shown in Figure 7. It is seen from this figure that the error probability is above 10^{-4} for $\gamma \geq 2$ and $\xi \leq 22$ dB. The single sample receiver can not combat the phase noise effectively for large γ, and the bit error rate severely degrades with increasing phase noise strength.

The error probability calculation is more involved for the multisample receivers. For the single-filter receiver, the error probability is given by (12) which can be written as

$$P_e = \frac{1}{2}\int_0^1 \exp(-\xi x/2M) q_{\gamma/M}(x)\, dx\ . \tag{47}$$

Using (39) with $\gamma_1 = \gamma$ and $\gamma_2 = \gamma/M$, we can rewrite (47) as

$$P_e = \frac{1}{2}\int_0^1 \exp(-\xi x/2M) M x^{M-1} q_\gamma(x^M)\, dx$$

which with the change of variable $y = x^M$ yields

$$P_e = \frac{1}{2}\int_0^1 \exp(-\xi y^{1/M}/2M) q_\gamma(y)\, dy\ . \tag{48}$$

We want to find the value of M that minimizes (48). Since no closed form for $q(\cdot)$ is available, this optimization involves evaluating this equation for $M = 1, 2, 3, \ldots$ until the optimum is reached[5]. Equation (48) has the desired form since the density function has the same parameter γ for all M.

The result of the computation of the bit error rate via (48) is shown in Figure 8. The plotted bit error rate values have been optimized over M. The optimal value

[5]In fact, we could search for optimal non-integer M for the single-filter receiver. This does not seem to be necessary since the observed transitions with integer values were smooth.

of M increases with SNR and the phase noise strength. This is because as the SNR increases the receiver prefers to use a smaller portion of the signal energy to reduce the effect of the phase noise with the same factor.

Finally we focus on the performance of the double filter multisampling receiver. In Section 2.2, we showed that the bit error rate for this receiver can be computed by first finding the M-fold convolution of the density function of $X(\gamma/M)$ and then removing the conditioning on r in Equation (19). However, the density function, $q_{\gamma/M}(x)$, changes with M, thus this method is computationally prohibitive. Therefore, we use the alternative approach which resulted in Equations (25) and (26). One can further reduce the complexity by expressing $\alpha(k)$ in terms of $q_\gamma(x)$ as follows:

$$\begin{aligned} \alpha(k) &= \int_0^1 x^k \exp(-\xi x/2M) q_{\gamma/M}(x)\, dx \\ &= \int_0^1 y^{k/M} \exp(-\xi y^{1/M}/2M) q_\gamma(y)\, dy\ . \end{aligned} \tag{49}$$

where we have used the same transformation used in obtaining (48).

Equation (25) in conjunction with (49) provides a method for the error probability computation. In performing the numerical computation the symmetry between the terms in the inner summation was exploited as outlined in Section 2.2. The resulting computation is a very efficient one in comparison with the direct convolutions. The results are shown in Figure 9. The bit error rates are again optimized with respect to M as in the single filter case. A comparison of the performance curves for the single sample receiver and multisample receiver with and without double filtering reveals that the increased receiver complexity by oversampling the signal and additional lowpass filtering may improve the error performance to a significant extent when the signal-to-noise ratio is high and the phase noise is strong. This can be seen from Figure 10 which shows the performance of the three receivers for $\gamma = 1$ and $\gamma = 4$. For low SNR values (up to 12 dB for $\gamma = 1$ and 9 dB for $\gamma = 4$), the best receiver is the simplest one, i.e. the optimal value for the multisampling receivers is $M = 1$. Only when the SNR is further increased, an increase in the receiver complexity is warranted. At a bit error rate of 10^{-9} with $\gamma = 1$, the double filter receiver has about 3.5 dB gain over multisample single-filter receiver and about 6 dB gain over the single sample receiver.

The error probability for the double filter receiver is very close to the lower bound obtained in Section 2.2, as illustrated by Figure 11 for $\gamma = 1$ and $\gamma = 4$. Therefore

the simple lower bound expression

$$P_e \simeq P_e\left(\xi\frac{4M}{\gamma}\left[1-\frac{2M}{\gamma}(1-e^{-\gamma/2M})\right]\right) \tag{50}$$

in conjunction with (19) provides a simple and reliable approximation to the probability of error. With the use of this closed-form result, the performance of a system with a given SNR and phase noise strength can be found by a straightforward optimization over M. It should be noted that the value of M that optimizes (50) is, in general, lower than the exact optimal value. This is due to the fact that the variation of r from its mean is neglected in the Jensen's bound that results in (50).

Since the phase noisy SNR r in (20) is the average of M i.i.d. random variables, a Gaussian approximation for large SNR and phase noise strength is also plausible. In fact, approximating the density of r with a Gaussian density on the interval $(0,\xi)$ yields an error probability which is very close to that shown in Figure 9. However, a closed-form result that is obtained by extending the range to the real line yields a poor estimate due to the large negative tail of $P_e(r)$.

The performance of the double-filter receiver as predicted by Figure 9 and approximated by Equation (50) is very close to that predicted by Foschini and his coworkers in [3]. This is because the truncated density tail of the linear approximation to the phase-noisy envelope is extremely low for a phase noise strength of γ/M, when M is around the optimal value. This is in contrast with the single-filter case where the agreement of our results with [3] is not as strong due to smaller values for the optimal value of M.

The signal-to-noise ratio needed to remain below a certain predetermined bit error rate increases with the introduction of phase noise. The amount of increase in the SNR is referred to as *phase noise penalty.* In Figure 12 the penalties for the three receivers are shown as a function of γ for a bit error rate of 10^{-9}. While the difference in the penalties for the single sample receiver and the multisample double-filter receiver is less than 1 dB for $\gamma \leq 0.5$, the latter receiver outperforms the former at higher phase noise strengths. This is in support of increased receiver complexity that is tailored towards combating the phase noise especially when $\gamma > 1$.

Advancing optical technology promises fiber-optic networks with data rates in excess of 100 Mbps per user and semiconductor lasers with linewidths of 5 MHz or less. With these parameters, the phase noise strength γ is in the range of 0.6 for which the penalty is about 1 dB with the double filter receiver and 2 dB with the single sample receiver. This means not only that the phase noise will not be a very

significant impairment on the bit error rate, but also that simple IF receivers will be sufficient for most future applications.

6 Extension to N-FSK

The results of the previous sections about binary FSK can be extended to FSK with N hypotheses (N-FSK) without much difficulty. The receiver structure is the same as Figure 2 with the number of branches increased from 2 to N. We present an analysis to the performance of N-FSK below. It must be emphasized that we hold the transmission rate of data bits and the energy per data bit constant while making comparisons among different N-ary FSK schemes, as in [20].

Let the possible hypotheses be $0, 1, \ldots, N-1$. Let the normalized filter outputs be V_i, $i = 0, 1, \ldots, N-1$, and assume a "0" is transmitted. Then V_i will be independent random variables with the density functions found from Section 2 as

$$p(v_0|r) = \left(\frac{v_0}{r}\right)^{(M-1)/2} e^{-(v_0+r)} I_{M-1}\left(\sqrt{4rv_0}\right) \quad v_0 \geq 0 \tag{51}$$

$$p(v_i) = \frac{v_i^{M-1}}{(M-1)!} e^{-v_i} \quad v_i \geq 0 \quad i = 1, 2, \ldots, N-1\,. \tag{52}$$

Let $P_N(r)$ denote the probability of a symbol error conditioned on the phase-noisy SNR r. Since a symbol error will occur when any of the V_i for $i \neq 0$ is greater than V_0, we have

$$\begin{aligned} P_N(r) &= 1 - \Pr\left(V_0 \geq \max\{V_1, \ldots, V_{N-1}\}|r\right) \\ &= 1 - \int \left[\Pr(V_1 \leq v_0)\right]^{N-1} p(v_0|r)\, dv_0 \end{aligned}$$

where we have used the fact that V_i for $i \geq 1$ are i.i.d. conditioned on r. Then with the use of (52), we obtain

$$P_N(r) = 1 - \int \left[1 - \sum_{k=0}^{M-1} \frac{v_0^k}{k!} e^{-v_0}\right]^{N-1} p(v_0|r)\, dv_0$$

which does not seem to be easily brought to a closed-form similar to (19). Instead of computing $P_N(r)$ numerically we will find upper and lower bounds in closed-form. We first use $1 - (1-y)^{N-1} \leq (N-1)y$ for $0 \leq y \leq 1$, to obtain

$$P_N(r) \leq (N-1) \int \sum_{k=0}^{M-1} \frac{v_0^k}{k!} e^{-v_0} p(v_0|r)\, dv_0 = (N-1) P_2(r) \tag{53}$$

where $P_2(r)$ is the conditional error probability for binary FSK given by (19). The upper bound in (53) is exactly the union bound. To obtain a lower bound to $P_N(r)$ we use

$$1-(1-y)^{N-1} \geq (N-1)y - \binom{N-1}{2} y^2$$

and we get

$$P_N(r) \geq (N-1)P_2(r) - \binom{N-1}{2} g_M(r)$$

where the function $g_M(r)$ is defined as

$$g_M(r) = \int \left(\sum_{k=0}^{M-1} \frac{v_0^k}{k!} e^{-v_0} \right)^2 p(v_0|r)\, dv_0 \,.$$

To obtain $g_M(r)$ in closed-form we first use

$$\begin{aligned} \left(\sum_{k=0}^{M-1} \frac{v_0^k}{k!} e^{-v_0} \right)^2 &= e^{-2v_0} \sum_{n=0}^{M-1} \sum_{m=0}^{M-1} \frac{v_0^{n+m}}{n!m!} \\ &= e^{-2v_0} \sum_{l=0}^{2(M-1)} \left[\frac{1}{l!} \sum_{k=0}^{\min(l,M-1)} \binom{l}{k} \right] v_0^l \end{aligned}$$

and we get

$$g_M(r) = \int e^{-2v_0} \sum_{l=0}^{2(M-1)} \frac{1}{l!} \beta_l v_0^l p(v_0|r)\, dv_0 \tag{54}$$

where we have defined

$$\beta_l \triangleq \sum_{k=0}^{\min(l,M-1)} \binom{l}{k} . \tag{55}$$

We proceed exactly as in Appendix A to evaluate the integral in (54) to obtain the result as

$$g_M(r) = \frac{e^{-2r/3}}{3^M} \sum_{n=0}^{2(M-1)} \frac{(r/3)^n}{n!} \sum_{l=n}^{2(M-1)} \binom{l+M-1}{l-n} \frac{\beta_l}{3^l} \tag{56}$$

which is quite similar to (19).

We finally have the bounds to the conditional error probability as

$$(N-1)P_2(r) - \binom{N-1}{2} g_M(r) \leq P_N(r) \leq (N-1)P_2(r) \tag{57}$$

in conjunction with Equations (19) and (56). Note that the two bounds are equal for $N = 2$. The bounds are tight for small N and large r. For $N \leq 16$ and $r \geq 14$ dB, the two bounds are within an order of magnitude for any M.

The symbol error probability P_N is obtained as the expected value of $P_N(r)$ optimized over M, i.e.

$$P_N = \min_M E_r \left[P_N(r) \right] .$$

Since only bounds to $P_N(r)$ are available in closed-form, we will perform the optimization for the upper bound and use the optimal value of M, say M^*, in the lower bound. The result is

$$(N-1)P_2 - \binom{N-1}{2} E_r \left[g_{M^*}(r) \right] \leq P_N \leq (N-1)P_2 \tag{58}$$

where P_2 is the binary FSK error probability found in the previous section. The bit error probability P_b is related to the symbol error probability as [12]

$$P_b = \frac{N/2}{N-1} P_N$$

which results in

$$\frac{N}{2}P_2 - \frac{N(N-2)}{4} E_r \left[g_{M^*}(r) \right] \leq P_b \leq \frac{N}{2} P_2 . \tag{59}$$

The upper bound to the actual bit error rate reveals that for N-FSK at a bit SNR value of ξ and a phase noise of γ is at most $N/2$ times the bit error rate of binary FSK at an SNR of $\xi \log_2 N$ and a phase noise strength of $\gamma \log_2 N$. The expectation in the lower bound can be computed without any convolutions, using the same method described in the previous section. Figure 13 shows the upper and lower bounds to P_b for $N = 4, 8, 16$ as well as the binary bit error rate for $\gamma = 1$. It is seen that the bounds are very close, and that the bit error rate improves with increasing N. At a bit error rate of 10^{-9} the improvement over binary FSK is 2.5 dB with $N = 4$, 4 dB with $N = 8$ and 4.8 dB with $N = 16$. Since the cost of the receiver and the bandwidth occupation are proportional to N, 4-FSK may be a desirable alternative to binary FSK for this value of phase-noise strength.

7 Conclusion

We have analyzed the performance of a binary FSK coherent optical system with an accurate treatment of the phase noise. We obtained closed-form expressions for the error probability conditioned on a normalized phase-noisy amplitude. These results can be used in conjunction with other approximations or solutions to the phase-noisy amplitude, e.g. the one given in [7]. In this work, we used an exponential

approximation to the square of the amplitude which was found to be accurate. We obtained the density of this approximation from its moments, which are available in closed-form, via a Legendre series method. As a result, we were able to obtain the performance of the receivers proposed in [3] with, we believe, more accuracy than in previous work. We also provided a simple and tight lower bound in closed-form. Our results show that the simple correlator receiver is adequate only for a certain range of signal-to-noise ratio and phase noise values, beyond which receivers which track the phase noise by oversampling improve considerably upon its performance.

We also extended the results to the case of N-FSK and obtained tight upper and lower bounds to the bit error rate. As a result, we observed that increasing the signal dimensionality may be desirable up to a certain limit to be robust against phase noise.

Appendix

A Derivation of Conditional Error Probability for Double–Filter Receiver

In this appendix, we derive (19) which gives the error probability of the multisample double-filter receiver condition on phase-noisy signal-to-noise ratio r. From (18) we have

$$\begin{aligned} P_e(r) &= \sum_{k=0}^{M-1} \frac{1}{k!} \int_0^\infty e^{-x} x^k \left(\frac{x}{r}\right)^{(M-1)/2} e^{-(x+r)} I_{M-1}\left(\sqrt{4rx}\right)\, dx \\ &= \frac{e^{-r}}{r^{(M-1)/2}} \sum_{k=0}^{M-1} \frac{1}{k!} A_k(r) \end{aligned} \tag{A.1}$$

where we have defined

$$A_k(r) \triangleq \int_0^\infty e^{-2x} x^{k+(M-1)/2} I_{M-1}\left(\sqrt{4rx}\right)\, dx\ .$$

With the change of variable $y = \sqrt{x}$ we obtain

$$A_k(r) = 2\int_0^\infty e^{-2y^2} y^{2k+M} I_{M-1}\left(\sqrt{4r}y\right)\, dy\ .$$

The integral above can be found from 11.4.28 of [15] in conjunction with $I_n(x) = j^{-n}J_n(jx)$ as

$$A_k(r) = \frac{r^{(M-1)/2}}{2^{k+M}} \frac{\Gamma(k+M)}{\Gamma(M)} F(k+M, M, r/2) \tag{A.2}$$

where F is the confluent hypergeometric function. Equation (A.2) in conjunction with (A.1) yields

$$P_e(r) = e^{-r} \sum_{k=0}^{M-1} \frac{1}{2^{k+M}k!} \frac{\Gamma(k+M)}{\Gamma(M)} F(k+M, M, r/2) \ . \tag{A.3}$$

Now we use the Kummer transformation (13.1.27) and the power series for a confluent hypergeometric function (13.1.2) from [15] to simplify (A.3) via

$$\begin{aligned} F(k+M, M, r/2) &= e^{r/2} F(-k, M, -r/2) \\ &= e^{r/2} \sum_{n=0}^{\infty} \frac{(-k)_n}{(M)_n} \frac{(-r/2)^n}{n!} \end{aligned} \tag{A.4}$$

where $(a)_n = a(a+1)\cdots(a+n-1)$. Since $(-k)_n = 0$ for $n > k$ the series above is a finite series. Furthermore, since $(-k)_n = (-1)^n k!/(k-n)!$ for $n \leq k$ and $(M)_n = (M+n-1)!/(M-1)!$, (A.4) becomes

$$F(k+M, M, r/2) = e^{r/2} \sum_{n=0}^{k} \frac{k!(M-1)!}{(k-n)!(M-1+n)!} \frac{(r/2)^n}{n!} \ . \tag{A.5}$$

Inserting (A.5) into (A.3), using $\Gamma(k) = (k-1)!$ and canceling the common terms, we obtain

$$P_e(r) = \frac{e^{-r/2}}{2^M} \sum_{k=0}^{M-1} \frac{1}{2^k} \sum_{n=0}^{k} \frac{(k+M-1)!}{(k-n)!(M-1+n)!} \frac{(r/2)^n}{n!} \tag{A.6}$$

which upon interchange of summations yields (19)

B Convexity of Conditional Error Probability

In this appendix we prove that the error probability conditioned on phase-noisy SNR, $P_e(r)$, is a convex function of r. From (19), we have

$$P_e(r) = \frac{e^{-r/2}}{2^M} \sum_{n=0}^{M-1} \frac{c_n(M-1)}{n!} \left(\frac{r}{2}\right)^n \tag{B.1}$$

where

$$c_n(M-1) = \sum_{k=n}^{M-1} \binom{k+M-1}{k-n} \frac{1}{2^k} \ . \tag{B.2}$$

Since the convexity of a function is not affected by multiplication by a positive constant and a scaling of its argument, it suffices to show the convexity of the function

$$h(r) = e^{-r}t(r)$$

where $t(r)$ is the polynomial of degree $M-1$ with coefficients $c_n(M-1)/n!$. The second derivative of $h(r)$ is

$$h''(r) = e^{-r}\left[t(r) - 2t'(r) + t''(r)\right]$$

where "prime"s denote derivatives. Thus, $P_e(r)$ is convex if and only if the polynomial

$$p(r) = t(r) - 2t'(r) + t''(r) \tag{B.3}$$

of degree $M-1$ is nonnegative for all $r \geq 0$. The latter will be satisfied if all the derivatives of $p(r)$ at $r = 0$ are nonnegative, by Taylor's theorem. The n'th derivative of $p(r)$ at $r = 0$ is given as

$$\begin{aligned} p^{(n)}(0) &= t^{(n)}(0) - 2t^{(n+1)}(0) + t^{(n+2)}(0) \\ &= c_n(M-1) - 2c_{n+1}(M-1) + c_{n+2}(M-1) \end{aligned}$$

Let $N = M-1$ and $d_n(N) = c_n(N) - c_{n+1}(N)$ for $n = 0, 1, \ldots, N$. Then we have

$$p^{(n)}(0) = d_n(N) - d_{n+1}(N)\,. \tag{B.4}$$

Thus, we need to show that the coefficients $d_n(N)$ are nonincreasing in n for every N. To do this, we need an intermediate result given in the following lemma.

Lemma B.1 The binomial coefficients satisfy

$$\sum_{k=0}^{j} 2^k \binom{J-k}{j-k} = \sum_{k=0}^{j} \binom{J+1}{k} \tag{B.5}$$

for $J = 0, 1, 2, \ldots$ and $j = 0, 1, \ldots, J$.

Proof : We will use induction on J. For $J = 0$, the claim is obviously true. For $J = 1$, the claim is easily verified for both $j = 0$ and $j = 1$. Now let's assume that the claim is true for J and for all $j \leq J$. We will show that this implies that the claim is also true for $J+1$ and for all $j \leq J+1$. Let $a_j(J)$ denote the left hand side of (B.5). For $0 \leq j \leq J$, we have

$$a_{j+1}(J+1) = \sum_{k=1}^{j+1} 2^k \binom{J+1-k}{j+1-k} + \binom{J+1}{j+1}$$

$$
\begin{aligned}
&= 2a_j(J) + \binom{J+1}{j+1} \\
&= 2\sum_{k=0}^{j}\binom{J+1}{k} + \binom{J+1}{j+1} \\
&= \binom{J+1}{0} + \sum_{k=0}^{j}\left[\binom{J+1}{k} + \binom{J+1}{k+1}\right] \\
&= \sum_{k=0}^{j+1}\binom{J+2}{k}
\end{aligned}
$$

which is the desired result. (The second equality above follows by a change in the index of summation, the third one follows by induction hypothesis, the fourth one is a simple rearrangement of terms and the last one follows by the well-known property of the Pascal triangle.) Finally $a_0(J+1) = 1$ trivially satisfies (B.5). Thus the claim is true for all $j \leq J+1$. This completes the proof by induction. □

A corollary to Lemma B.1 provides a closed-form to $d_n(N)$.

Corollary The difference coefficients $d_n(N)$ satisfy

$$
d_n(N) = \frac{1}{2^N}\binom{2N+1}{N-n} \quad n = 0, 1, \ldots, N. \tag{B.6}
$$

Proof : From (B.2) we have

$$
\begin{aligned}
2^N c_n(N) &= \sum_{k=n}^{N} 2^{N-k}\binom{k+N}{k-n} \\
&= \sum_{l=0}^{N-n} 2^l \binom{2N-l}{N-n-l}
\end{aligned}
$$

where we have used a change of index $l = N - k$. Now we use Lemma B.1 with $J = 2N$ and $j = N - n$ to get

$$
c_n(N) = \frac{1}{2^N}\sum_{k=0}^{N-n}\binom{2N+1}{k}
$$

which results in

$$
\begin{aligned}
d_n(N) &= c_n(N) - c_{n+1}(N) \\
&= \frac{1}{2^N}\binom{2N+1}{N-n}
\end{aligned}
$$

as claimed. □

As a result of this corollary, we see that for a fixed N, $d_n(N)$ are decreasing with n for $0 \leq n \leq N$. This means that the polynomial $p(r)$ is nonnegative for $r \geq 0$, and consequently that the conditional error probability $P_e(r)$ is convex for any M, as explained before.

An implication of the convexity of $P_e(r)$ is that the conditional error probability can not be improved by time-sharing between two phase-noisy SNRs r_1 and r_2, since $P_e(\lambda r_1 + (1-\lambda)r_2) \leq \lambda P_e(r_1) + (1-\lambda)P_e(r_2)$ for all $0 \leq \lambda \leq 1$.

C Tail Behavior of $q(x)$

In this appendix we concentrate on the decay properties of the density function $q(x)$ around its $x = 1$ tail. We prove that this decay is exponential as described in the following property.

Property: The density function $q(x)$ has an exponential decay around $x = 1$, characterized as

$$\lim_{x \to 1^-} (1-x) \ln q(x) = -\gamma/8 \; . \tag{C.1}$$

To prove this property we need two lemmas. We first define a transformation which is closely related to Mellin transform [21].

Definition: The *moment transform* of a function $f(x)$ defined on $(0,1)$ is given by

$$F(t) = \int_0^1 x^t f(x)\, dx \tag{C.2}$$

where $t \geq 0$.

It can be observed that for any bounded, continuous function the moment transform exists, and that it vanishes as $t \to \infty$. The behavior of a function in its $x = 1$ tail is related to the asymptotic behavior of its moment transform. The following lemma formalizes this notion.

Lemma C.1 Let f and g be two bounded, analytic functions on $(0,1)$ whose moment transforms satisfy $F(t) \leq G(t)$ for all $t \geq t_0$. Then there exists $\epsilon > 0$ such that $f(x) \leq g(x)$ for all $x > 1 - \epsilon$.

Proof : Let $h = g - f$. h is bounded and analytic with moment transform $H = G - F$. Since h is analytic, there is a neighborhood of 1 in which h does not change sign, say $(1-\delta, 1)$ for some $\delta > 0$. Let

$$\begin{aligned} A &= \max\{|h(x)| : \; x \in (0,1)\} \\ B &= \int_{1-\delta}^1 h(x)\, dx \; . \end{aligned}$$

Suppose $h(x) \leq 0$ on $(1-\delta, 1)$. Then the following inequalities hold,

$$\begin{aligned} \int_0^{1-\delta} x^t h(x)\, dx &\leq \int_0^{1-\delta} x^t |h(x)|\, dx \leq A\frac{(1-\delta)^{t+1}}{t+1} \\ \int_{1-\delta}^{1} x^t h(x)\, dx &\leq (1-\delta)^t B\ , \end{aligned}$$

which implies

$$H(t) \leq (1-\delta)^t \left[\frac{A(1-\delta)}{t+1} + B\right]\ . \tag{C.3}$$

However, since $A > 0$ and $B < 0$, t can be made large enough to make the right hand side of (C.3) negative. This is a contradiction, since $H(t) \geq 0$ for all $t \geq t_0$. The proof is complete with the choice of an ϵ with $0 < \epsilon < \delta$. □

The following lemma establishes a relation between the moment transform of a function and the Laplace transform of its $x = 1$ tail.

Lemma C.2 Let f be a bounded function on $(0,1)$ and let f_ϵ be defined as

$$f_\epsilon(x) = \begin{cases} f(1-x) & 0 \leq x < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

for $0 < \epsilon < 1$. If $Q_\epsilon(t)$ is the Laplace transform of $f_\epsilon(x)$ and $F(t)$ is the moment transform of $f(x)$, then $Q_\epsilon(t) \simeq F(t)$ for large t and small ϵ.

Proof : $Q_\epsilon(t)$ is defined as

$$Q_\epsilon(t) = \int_0^\epsilon f(1-x)e^{-tx}\, dx\ .$$

With small ϵ we use $e^{-tx} \simeq (1-x)^t$ for $0 < x < \epsilon$, to obtain

$$Q_\epsilon(t) \simeq \int_0^\epsilon f(1-x)(1-x)^t\, dx = \int_{1-\epsilon}^1 f(x)x^t\, dx\ .$$

Since x^t becomes smaller in the interval $(0, 1-\epsilon)$ with increasing t, the rightmost integral will be arbitrarily close to $F(t)$ for large t. Therefore, the quantity $|Q_\epsilon(t) - F(t)|$ can be made arbitrarily close to 0 for sufficiently small ϵ and large t. □

With these two lemmas at hand we now proceed with the proof of the property stated at the beginning of this appendix.

Proof of Property : Lemma C.2 states that the tail of the density function $q(x)$ around $x = 1$ can be found as the inverse Laplace transform of its moment transform for large arguments. The moment transform, $\mu(t)$, of $q(x)$ is given by (33) and can be written as

$$\mu(t) = \left(\frac{2\sqrt{2\gamma t}}{1 - e^{-\sqrt{2\gamma t}}}\right)^{1/2} e^{-\sqrt{2\gamma t}/2}\ .$$

For large t, the first factor is bounded between 1 and $\sqrt{t}$. Thus, for large t we have

$$e^{-\sqrt{2\gamma t}/2} \leq \mu(t) \leq \sqrt{t} e^{-\sqrt{2\gamma t}/2} . \tag{C.4}$$

Then from the two lemmas we proved, the function $q(1-x)$ for small x will be bounded by the two functions whose Laplace transforms yield the bounds in (C.4). With the use of Formula 29.3.82 of [15] we obtain

$$\sqrt{\frac{\gamma}{8\pi y^3}} e^{-\gamma/8y} \leq q(1-y) \leq \sqrt{\frac{\gamma}{8\pi y^3}} \left(\frac{\gamma}{4y} - 1 \right) e^{-\gamma/8y}$$

for small y, which upon substitution of x for $1-y$ becomes

$$\sqrt{\frac{\gamma}{8\pi(1-x)^3}} e^{-\gamma/8(1-x)} \leq q(x) \leq \sqrt{\frac{\gamma}{8\pi(1-x)^3}} \left(\frac{\gamma}{4(1-x)} - 1 \right) e^{-\gamma/8(1-x)} \tag{C.5}$$

for x close to 1. We observe that the dominant tail behavior of $q(x)$ is exponential in $(-\gamma/8(1-x))$. Equivalently, by taking logarithms of all sides in (C.5), we obtain

$$\lim_{x \to 1^-} (1-x) \ln q(x) = -\frac{\gamma}{8} \tag{C.6}$$

where we have used $\lim_{x \to 0} x \ln x = 0$. This completes the proof. □

List of Figures

References

[1] C. H. Henry, "Theory of linewidth of semiconductor lasers," *IEEE Journal of Quantum Electronics*, vol. QE-18, pp. 259–264, February 1982.

[2] J. Salz, "Coherent lightwave communications," *AT&T Technical Journal*, vol. 64, pp. 2153–2209, December 1985.

[3] G. J. Foschini, L. J. Greenstein, and G. Vannucci, "Noncoherent detection of coherent lightwave signals corrupted by phase noise," *IEEE Transactions on Communications*, vol. COM-36, pp. 306–314, March 1988.

[4] G. J. Foschini, G. Vannucci, and L. J. Greenstein, "Envelope statistics for filtered optical signals corrupted by phase noise," *IEEE Transactions on Communications*, vol. COM-37, pp. 1293–1302, December 1989.

[5] L. G. Kazovsky, "Impact of laser phase noise on optical heterodyne communications systems," *Journal of Optical Communications*, vol. 7, pp. 66–77, June 1986.

[6] I. Garrett and G. Jacobsen, "Phase noise in weakly coherent systems," *IEE Proceedings*, vol. 136, pp. 159–165, June 1989.

[7] I. Garrett, D. J. Bond, J. B. Waite, D. S. L. Lettis, and G. Jacobsen, "Impact of phase noise in weakly coherent systems: A new and accurate approach," *Journal of Lightwave Technology*, vol. 8, pp. 329–337, March 1990.

[8] L. G. Kazovsky and O. K. Tonguz, "ASK and FSK coherent lightwave systems: A simplified approximate analysis," *Journal of Lightwave Technology*, vol. 8, pp. 338–352, March 1990.

[9] G. J. Foschini and G. Vannucci, "Characterizing filtered light waves corrupted by phase noise," *IEEE Transactions on Information Theory*, vol. IT-34, pp. 1437–1448, November 1988.

[10] L. G. Kazovsky, P. Meissner, and E. Patzak, "ASK multiport optical homodyne systems," *Journal of Lightwave Technology*, vol. 5, pp. 770–791, June 1987.

[11] W. C. Lindsey, "Error probabilities for rician fading multichannel reception of binary and N-ary signals," *IEEE Transactions on Information Theory*, vol. IT-10, pp. 339–350, October 1964.

[12] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 1983.

[13] J. R. Barry and E. A. Lee, "Performance of coherent optical receivers," *Proceedings of the IEEE*, vol. 78, pp. 1369–1394, August 1990.

[14] S. M. Ross, *Stochastic Processes.* New York: John Wiley & Sons, 1983.

[15] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions.* New York: Dover, 1972.

[16] R. G. Gallager, *Information Theory and Reliable Communication.* New York: John Wiley & Sons, 1968.

[17] H. L. Van Trees, *Detection, Estimation, and Modulation Theory.* New York: John Wiley & Sons, 1968.

[18] A. Papoulis, *Probability, Random Variables and Stochastic Processes.* New York: McGraw-Hill, second ed., 1984.

[19] R. V. Churchill and J. W. Brown, *Fourier Series and Boundary Value Problems.* New York: McGraw-Hill, third ed., 1978.

[20] L. J. Greenstein, G. Vannucci, and G. J. Foschini, "Optical power requirements for detecting OOK and FSK signals corrupted by phase noise," *IEEE Transactions on Communications*, vol. COM-37, pp. 405–407, April 1989.

[21] G. Arfken, *Mathematical Methods for Physicists.* Academic Press, 1985.

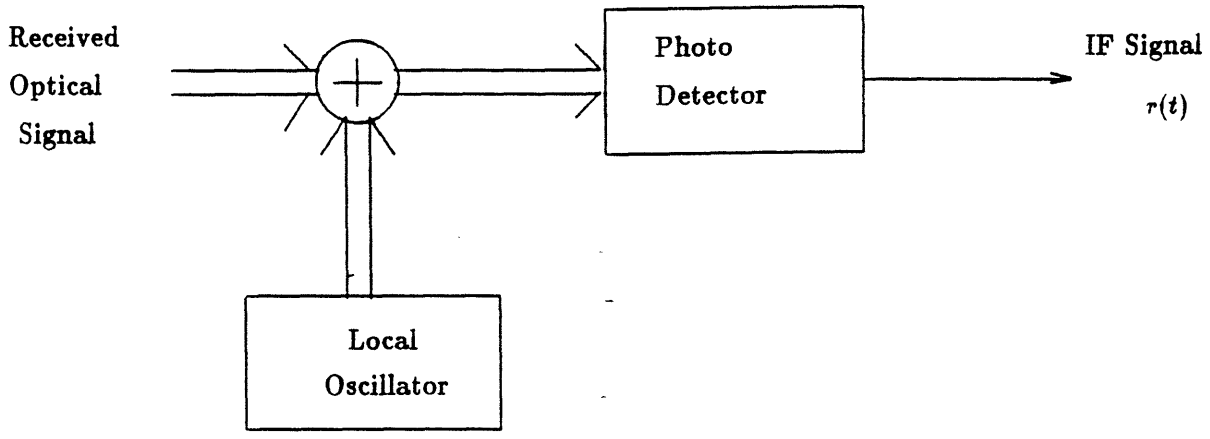
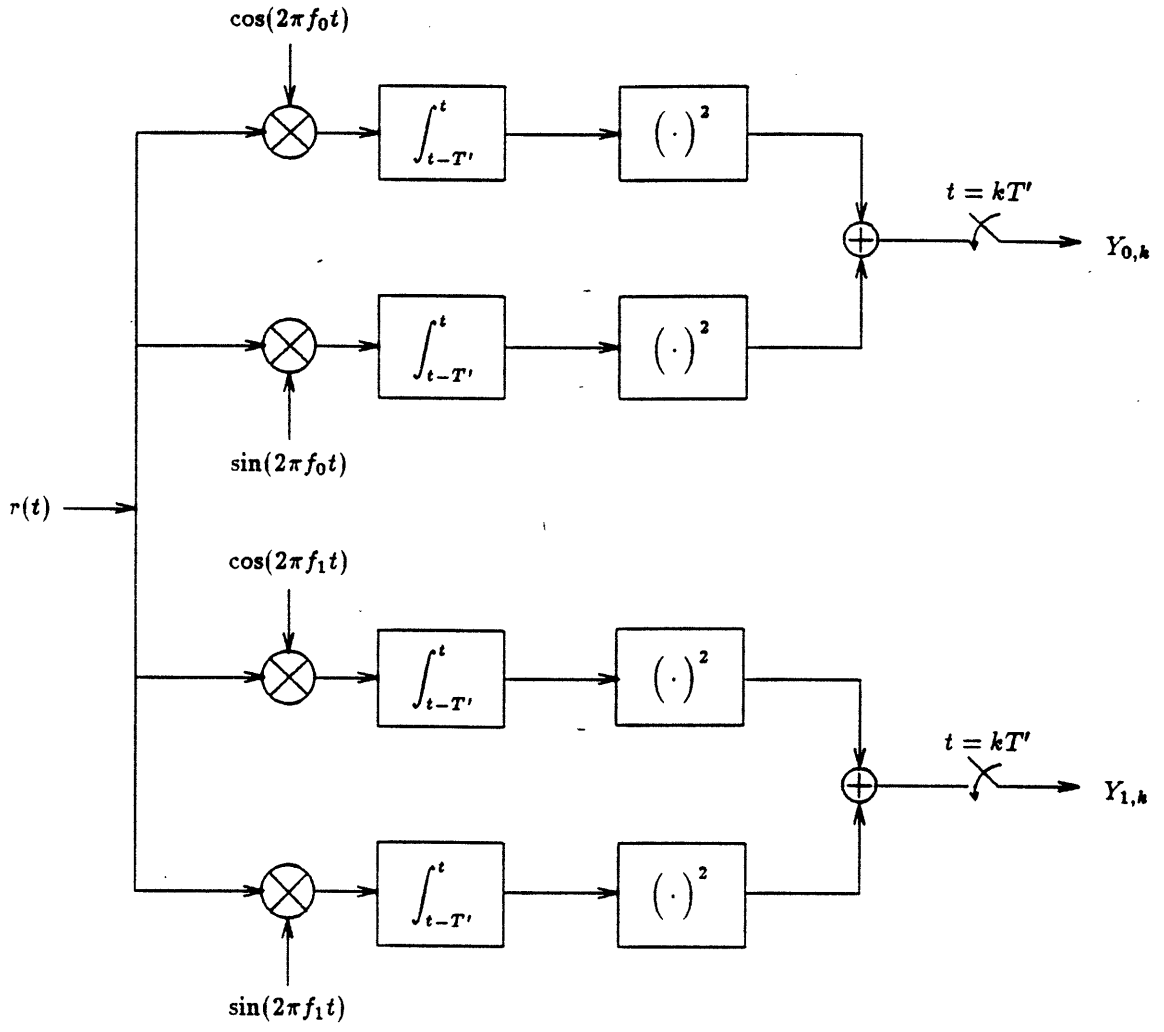
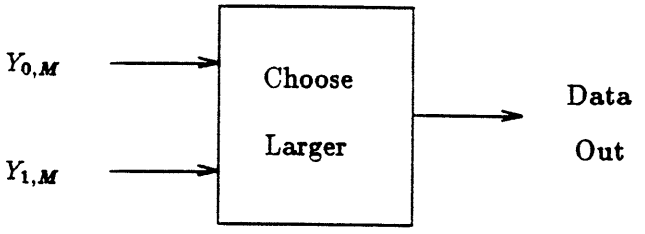

Received
Optical
Signal
Photo
Detector
IF Signal
$r(t)$
Local
Oscillator

Fig. 1

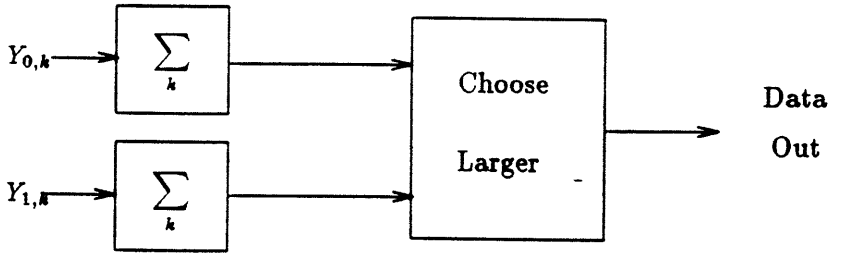

$\cos(2\pi f_0 t)$
$\int_{t-T'}^{t}$
$(\cdot)^2$
$t = kT'$
$Y_{0,k}$
$\sin(2\pi f_0 t)$
$r(t)$
$\cos(2\pi f_1 t)$
$t = kT'$
$Y_{1,k}$
$\sin(2\pi f_1 t)$

a)

Fig. 2 a)


$Y_{0,M}$
$Y_{1,M}$
Choose
Larger
Data
Out

b)


$Y_{0,k}$
$Y_{1,k}$
$\sum_k$
$\sum_k$
Choose
Larger
Data
Out

c)

Fig. 2b), c)

Fig. 3

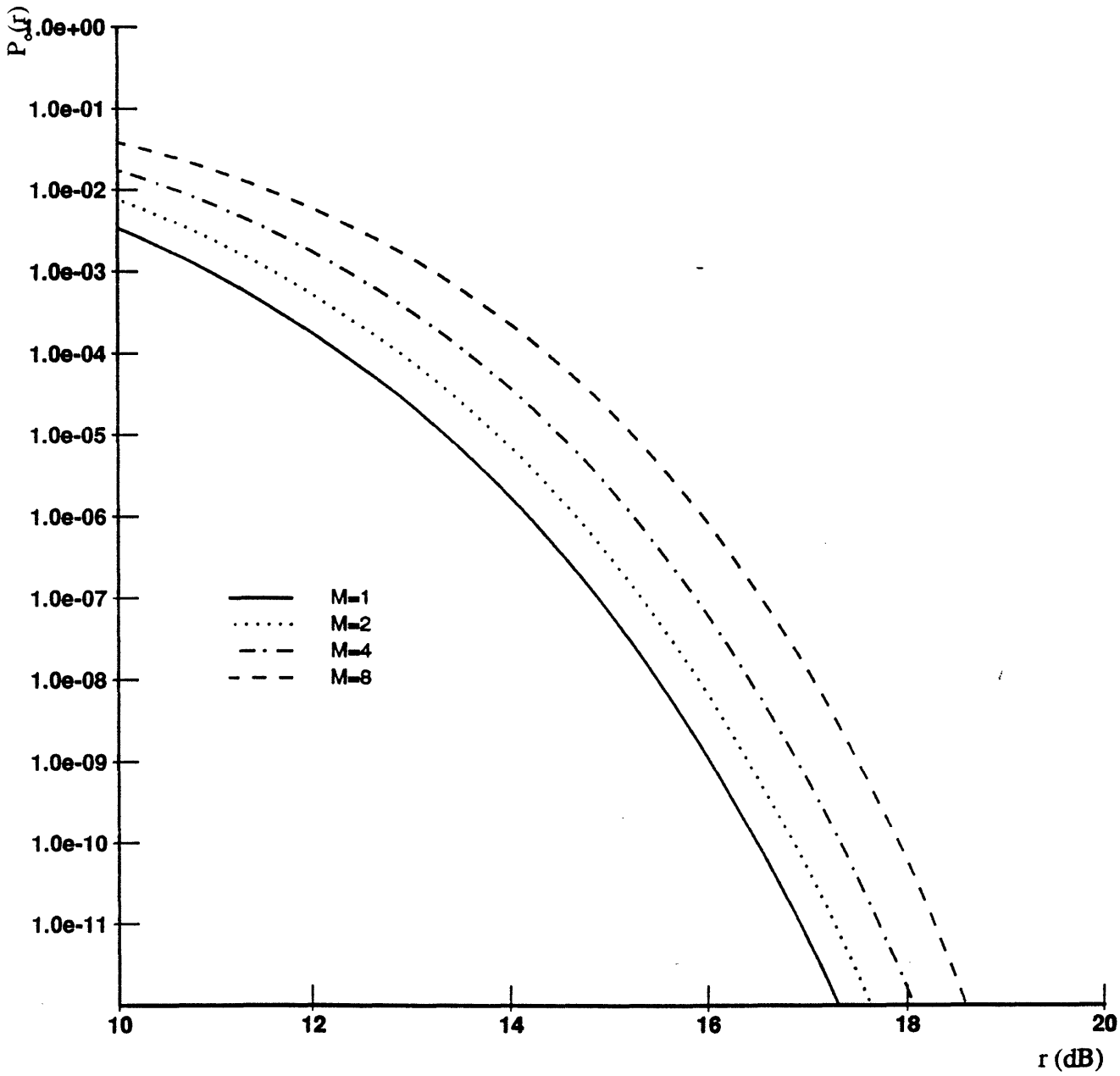

P_d(r)
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
M=1
M=2
M=4
M=8
10
12
14
16
18
20
r (dB)

Fig. 4

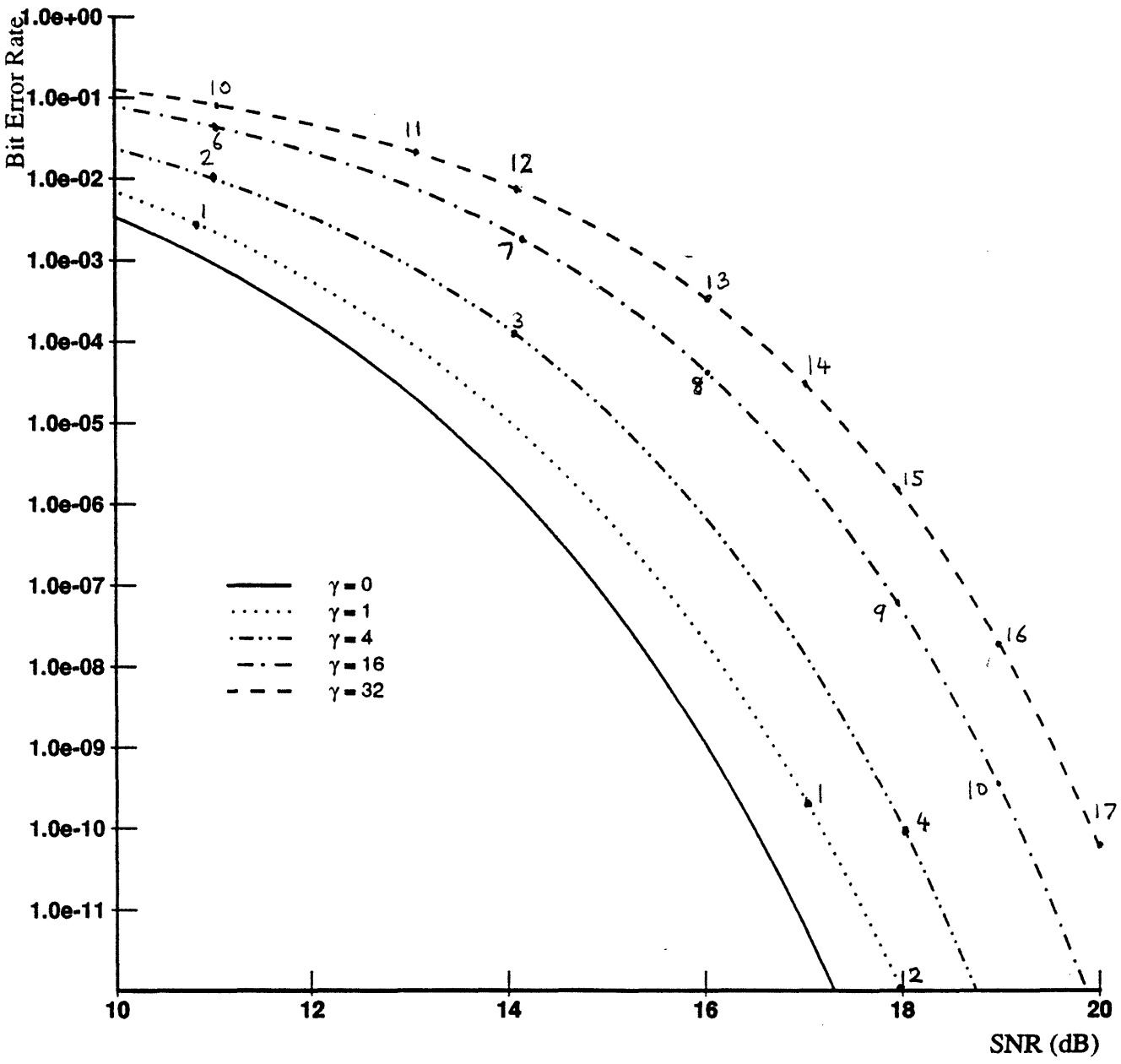

Bit Error Rate
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
10
12
14
16
18
20
SNR (dB)
γ = 0
γ = 1
γ = 4
γ = 16
γ = 32

Fig. 5a)

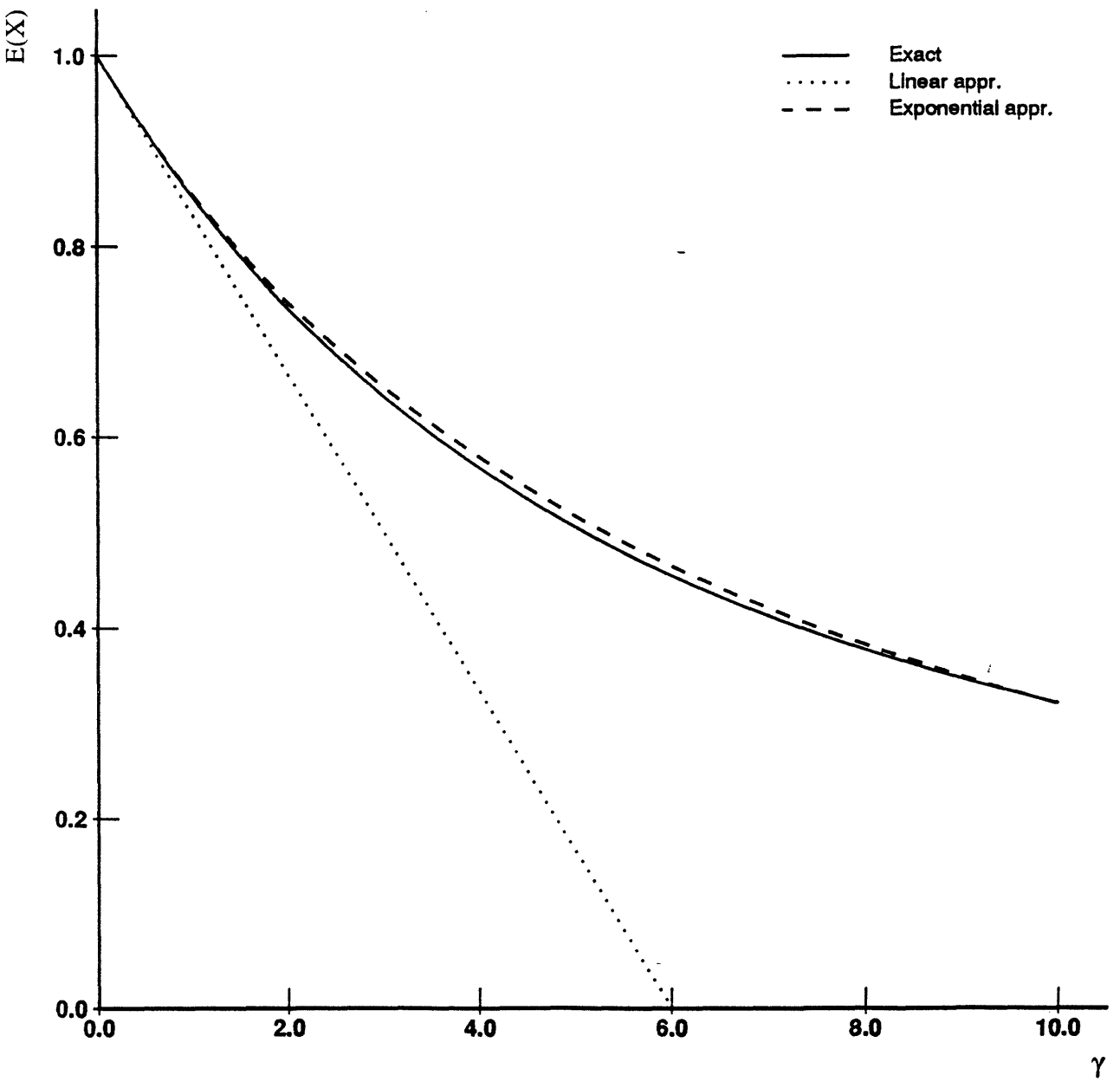

E(X)
1.0
0.8
0.6
0.4
0.2
0.0
0.0
2.0
4.0
6.0
8.0
10.0
γ
Exact
Linear appr.
Exponential appr.

Fig. 5 b)

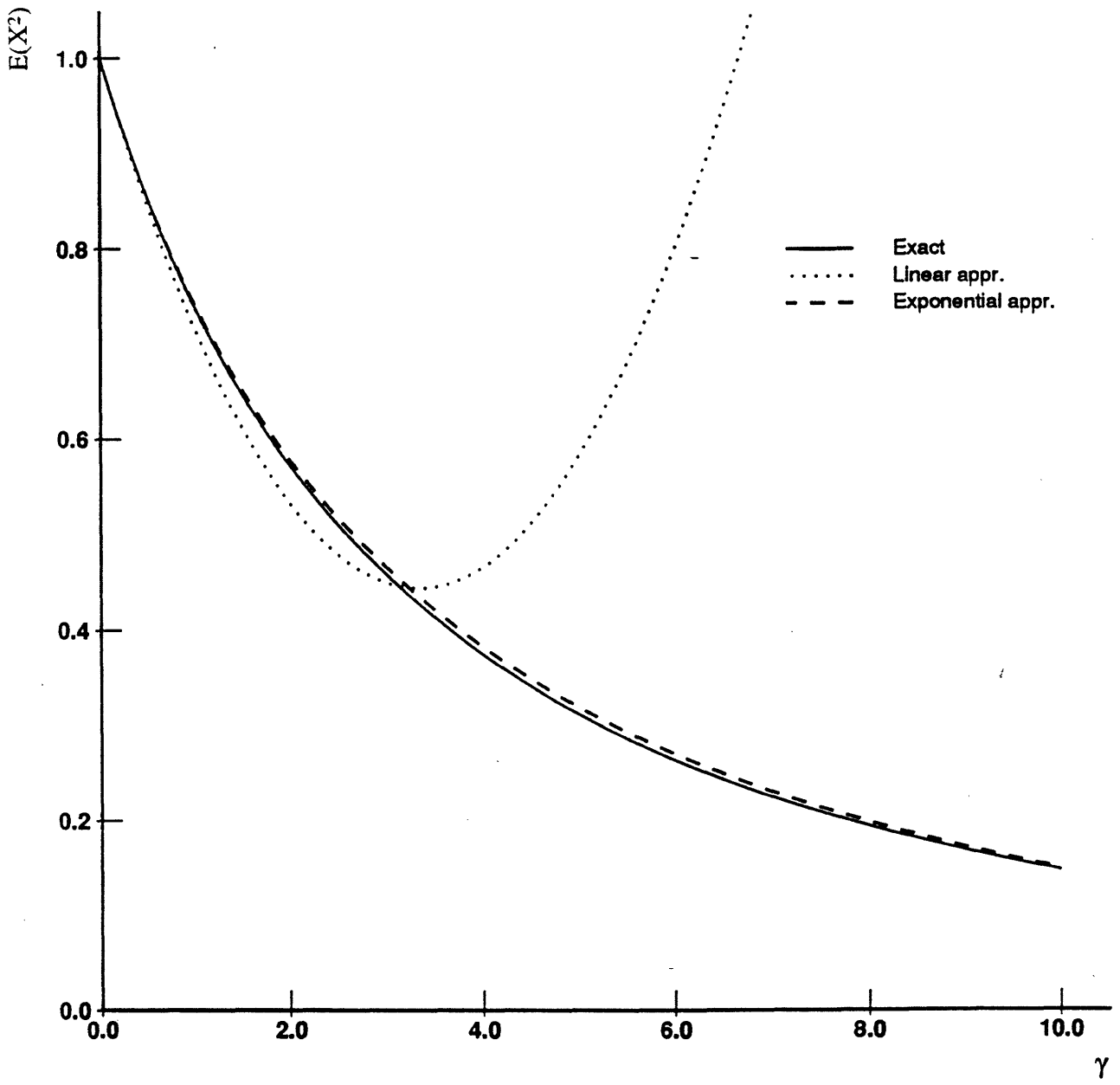

E(X²)
1.0
0.8
0.6
0.4
0.2
0.0
0.0
2.0
4.0
6.0
8.0
10.0
γ
Exact
Linear appr.
Exponential appr.

Fig. 6a)

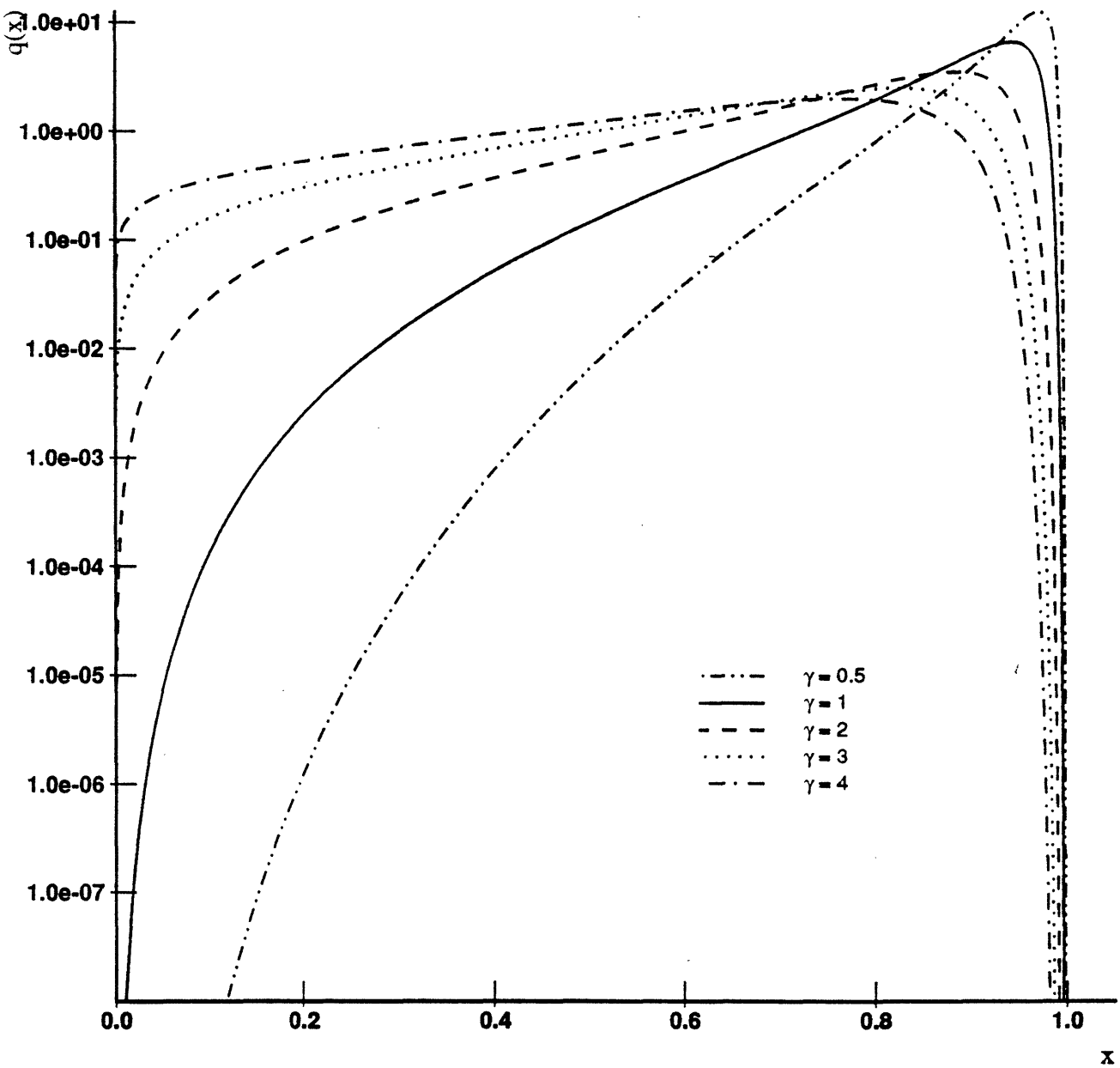

g(x)
1.0e+01
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
0.0
0.2
0.4
0.6
0.8
1.0
x
γ = 0.5
γ = 1
γ = 2
γ = 3
γ = 4

Fig. 6b)

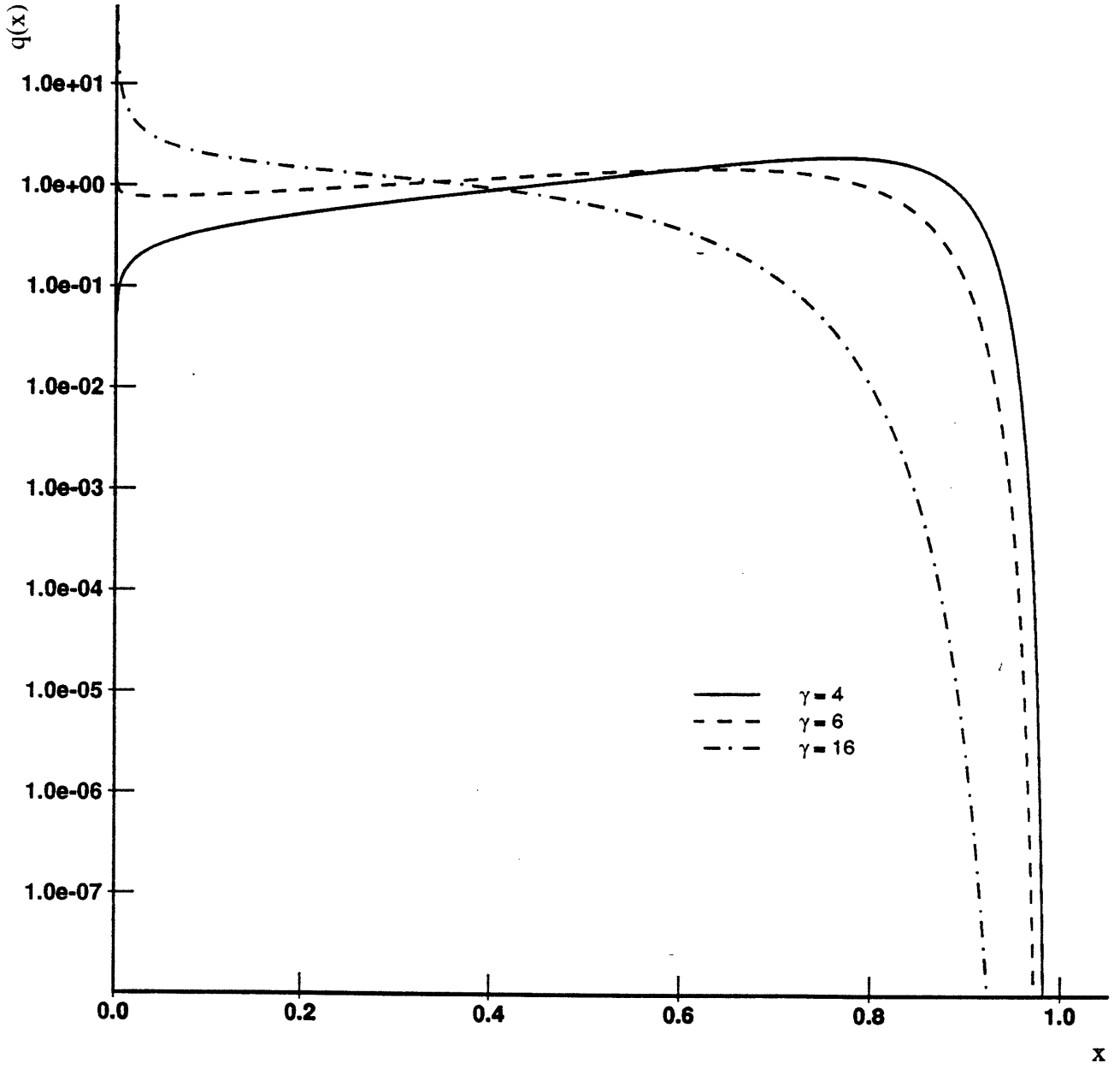

q(x)
1.0e+01
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
0.0
0.2
0.4
0.6
0.8
1.0
x
γ = 4
γ = 6
γ = 16

Fig. 7

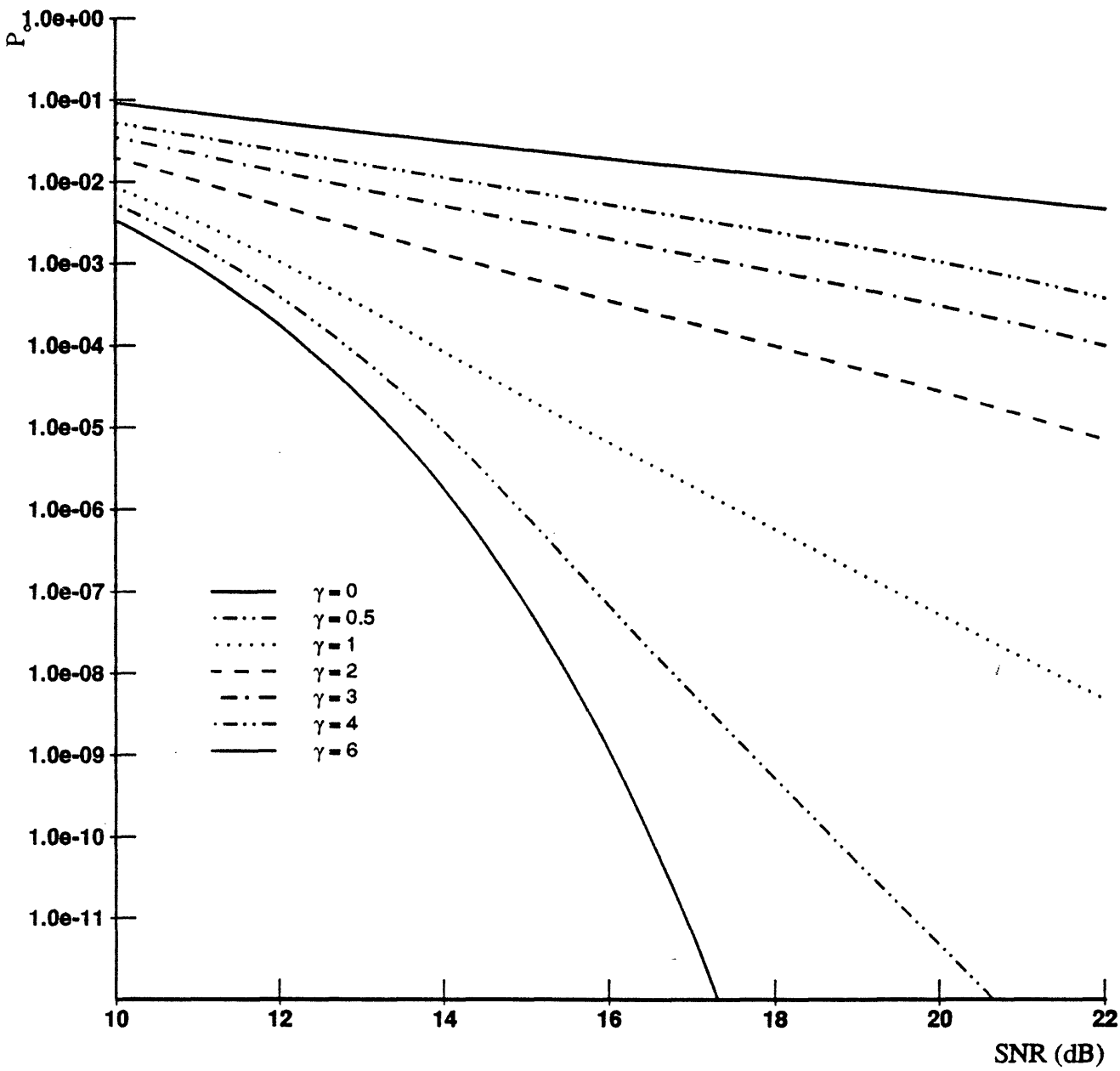

P_d
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
γ = 0
γ = 0.5
γ = 1
γ = 2
γ = 3
γ = 4
γ = 6
10
12
14
16
18
20
22
SNR (dB)

Fig. 8

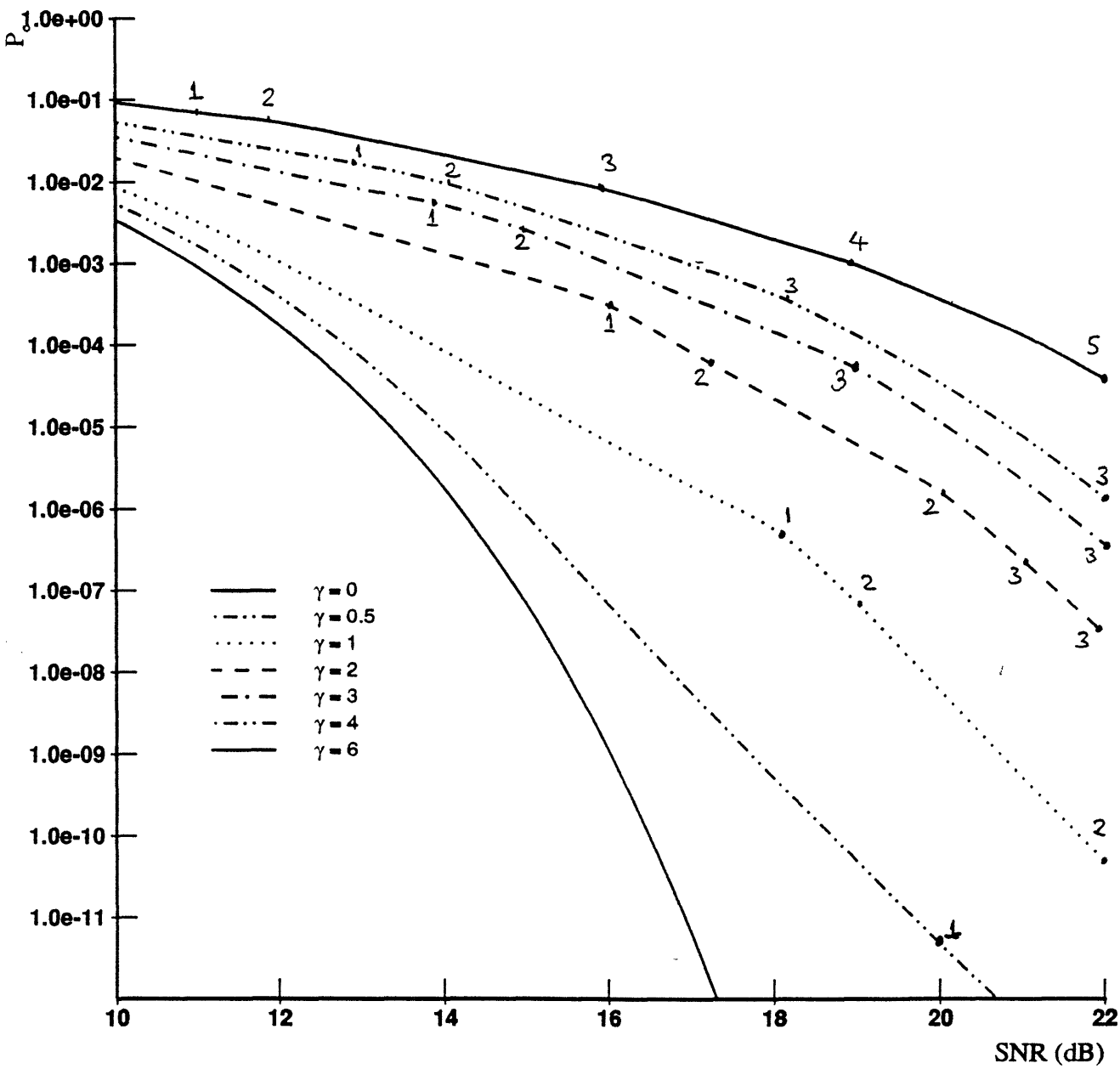

P
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
γ = 0
γ = 0.5
γ = 1
γ = 2
γ = 3
γ = 4
γ = 6
10
12
14
16
18
20
22
SNR (dB)

Fig. 9

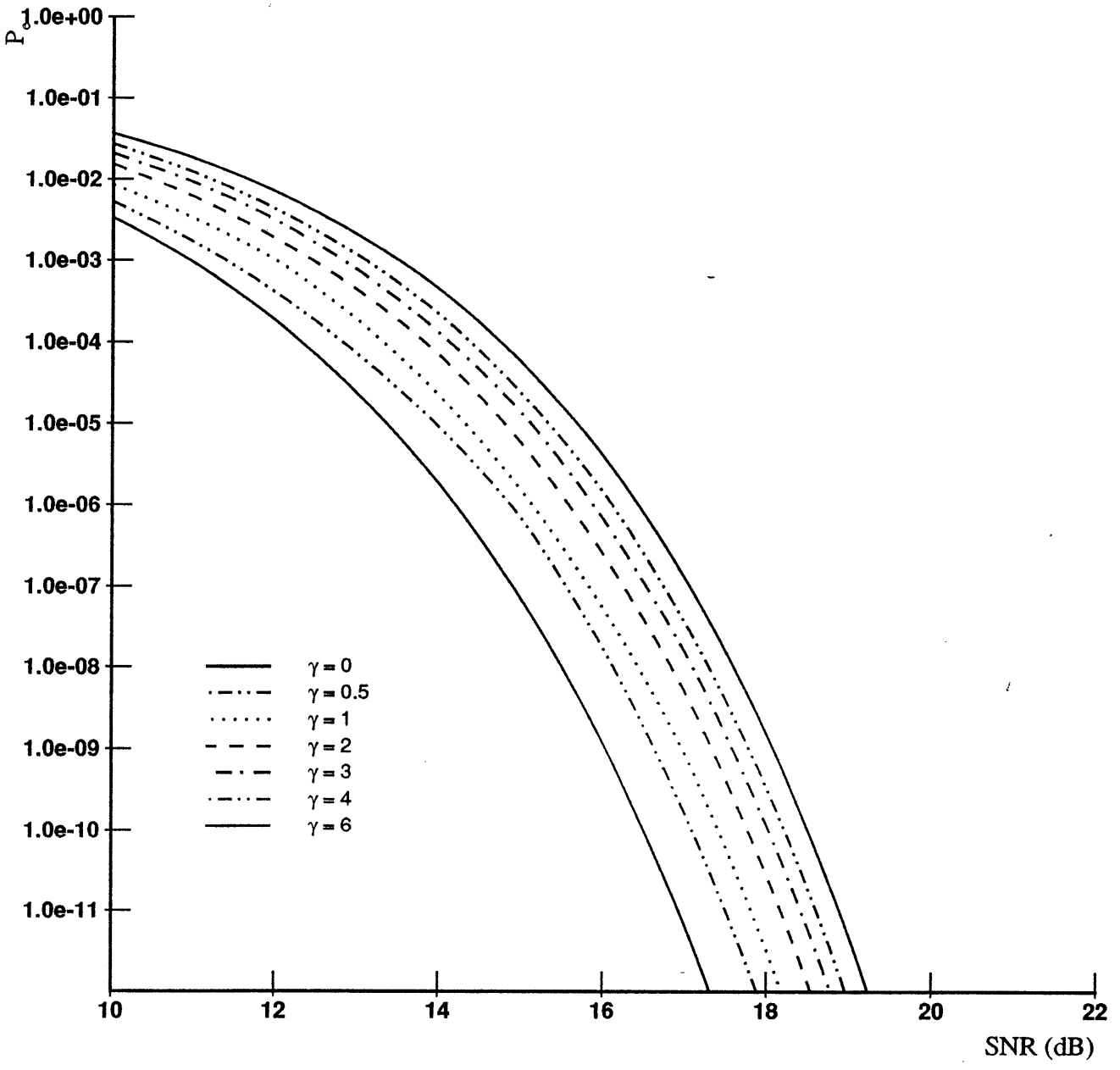

P_d
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
γ = 0
γ = 0.5
γ = 1
γ = 2
γ = 3
γ = 4
γ = 6
10
12
14
16
18
20
22
SNR (dB)

Fig. 10 a)

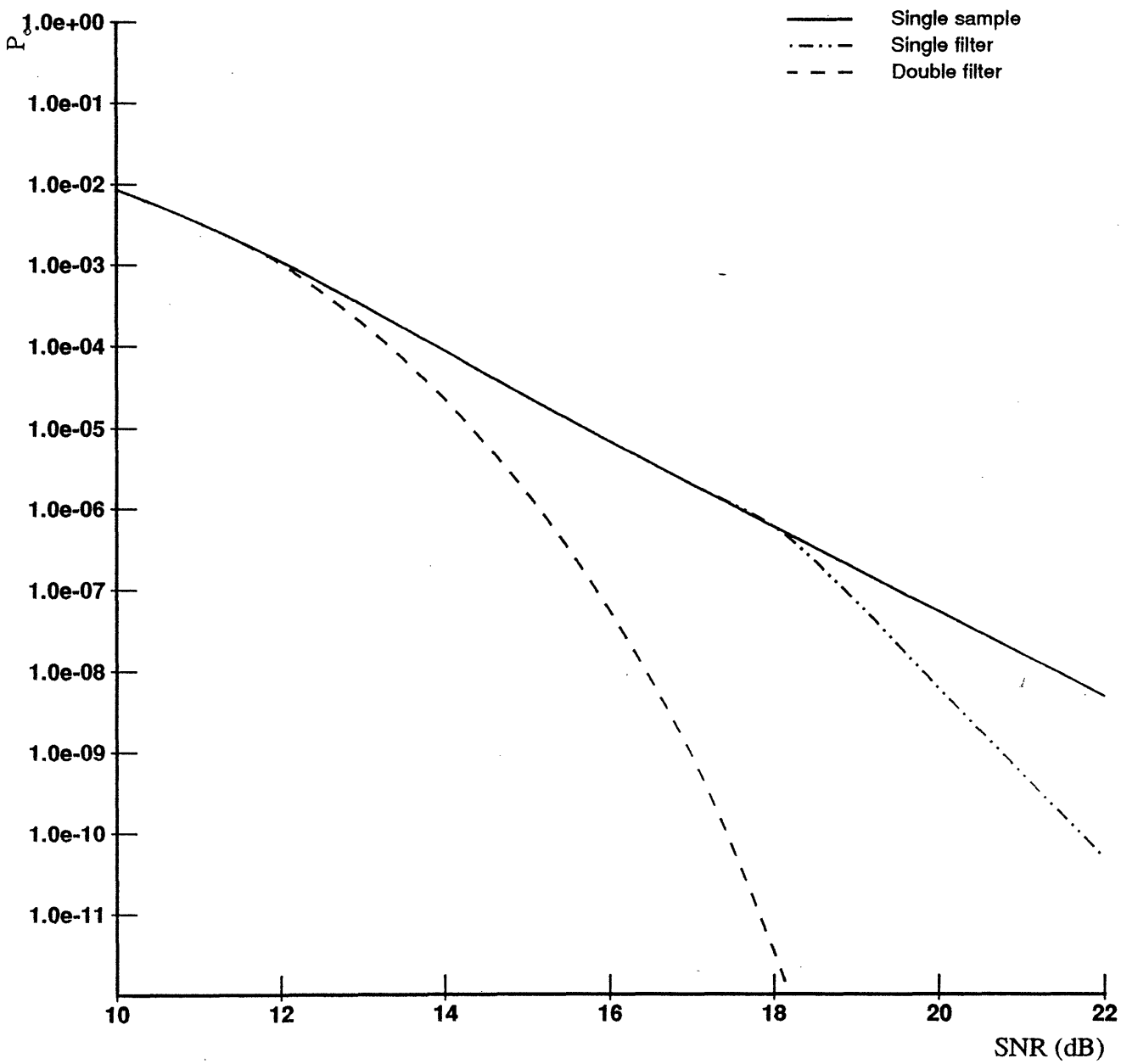

Single sample
Single filter
Double filter
P
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
10
12
14
16
18
20
22
SNR (dB)

Fig. 10 b)

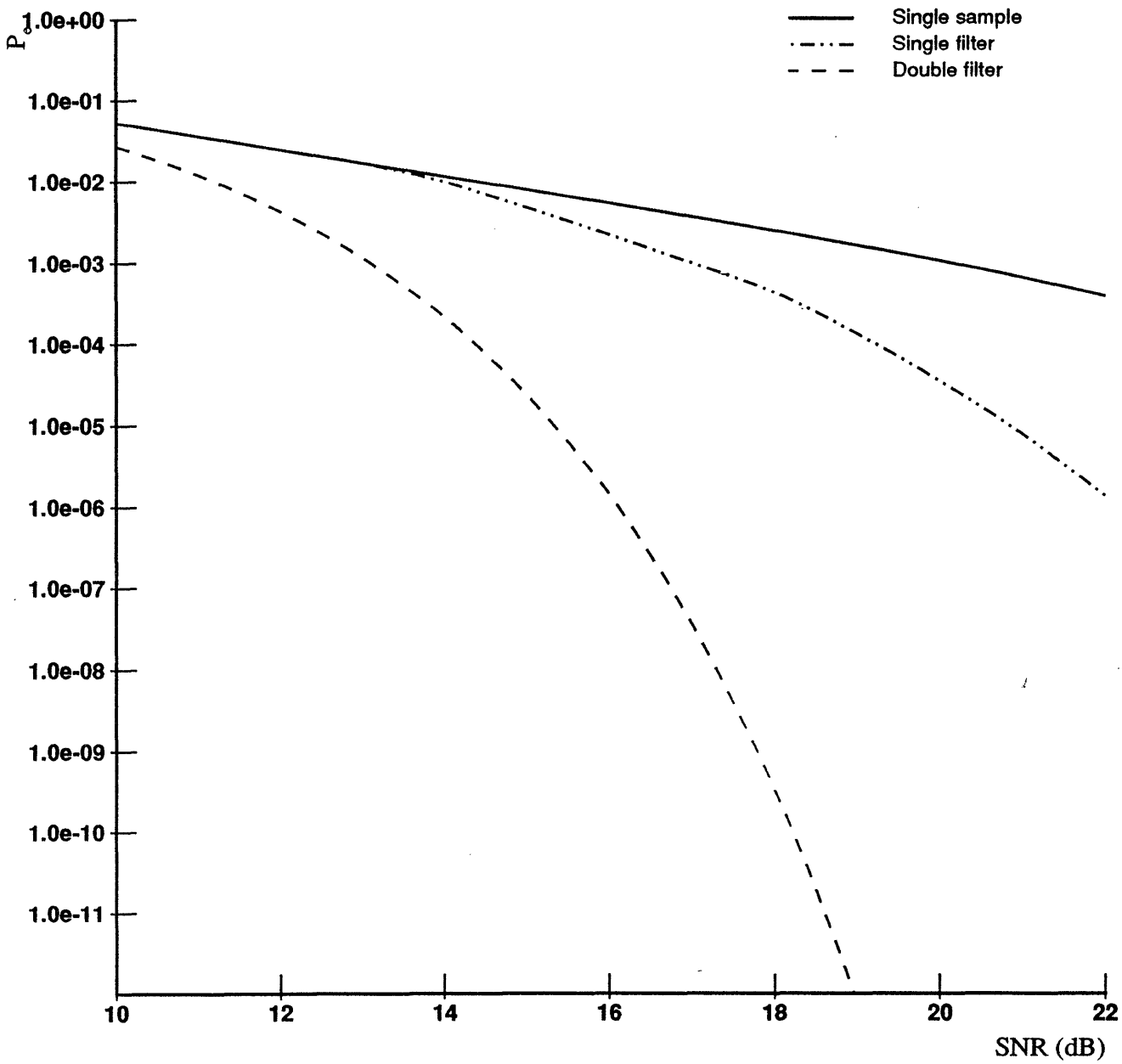

Single sample
Single filter
Double filter
P_d
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
10
12
14
16
18
20
22
SNR (dB)

Fig. 11

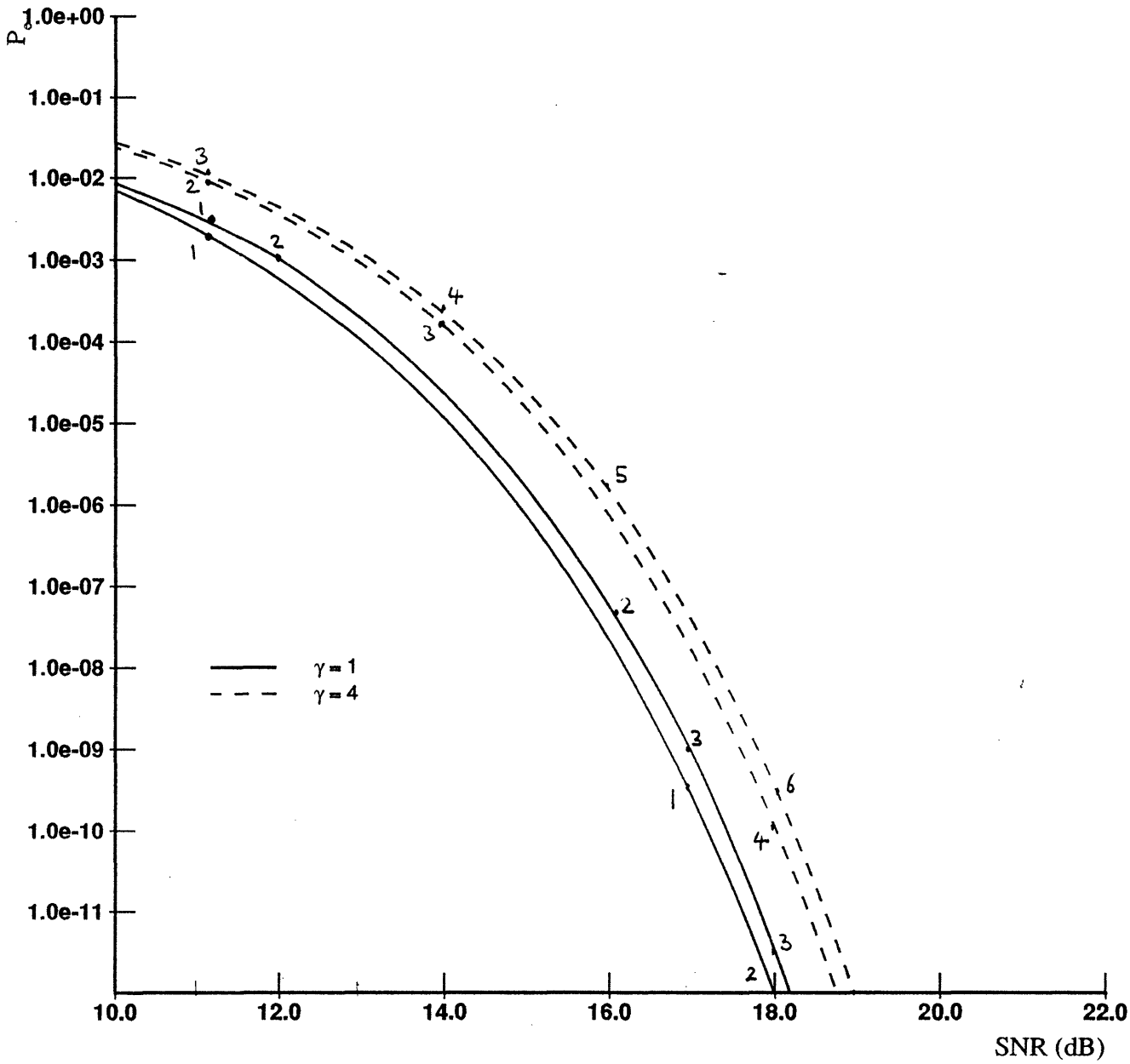

P
1.0e+00
1.0e-01
1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
γ = 1
γ = 4
10.0
12.0
14.0
16.0
18.0
20.0
22.0
SNR (dB)

Fig. 12

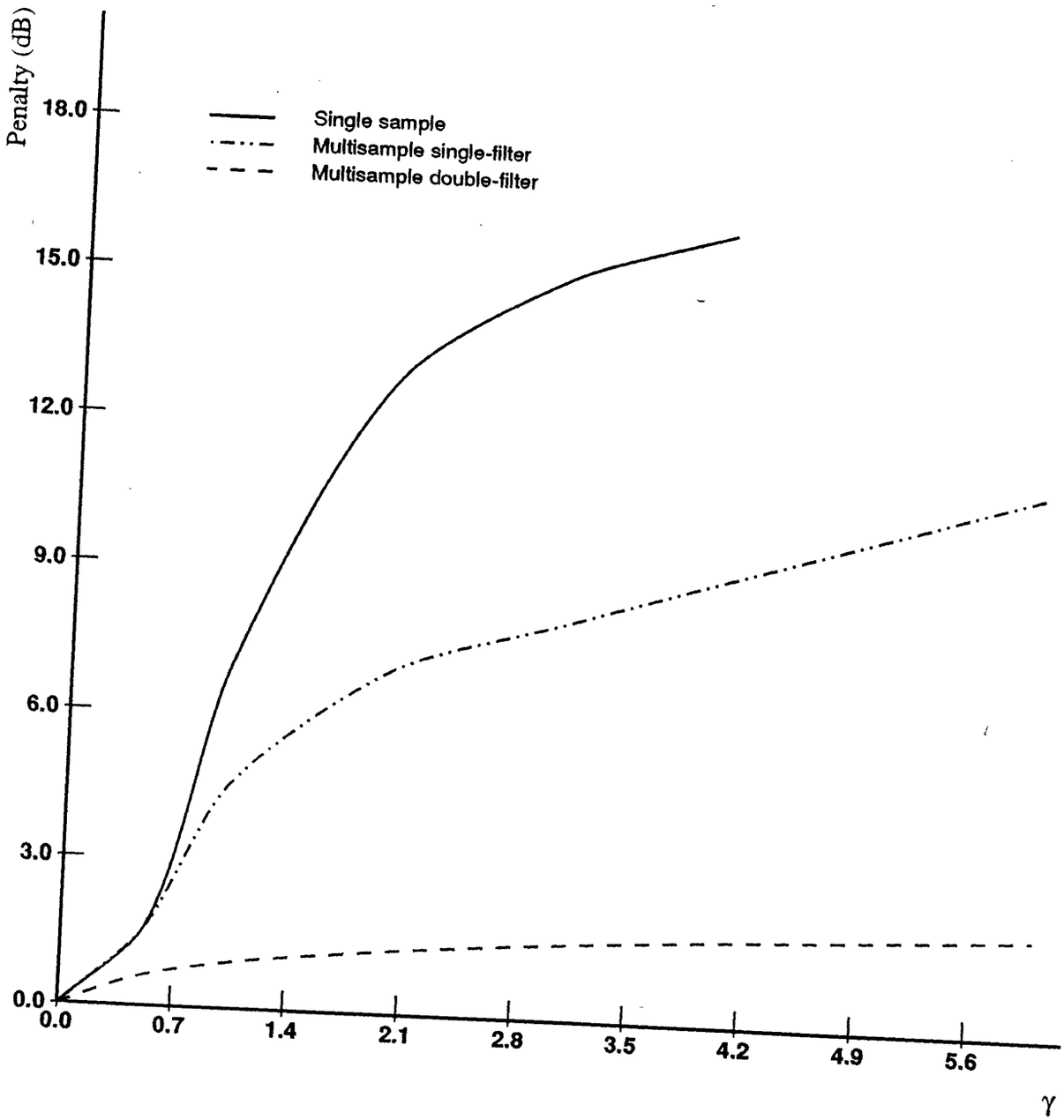
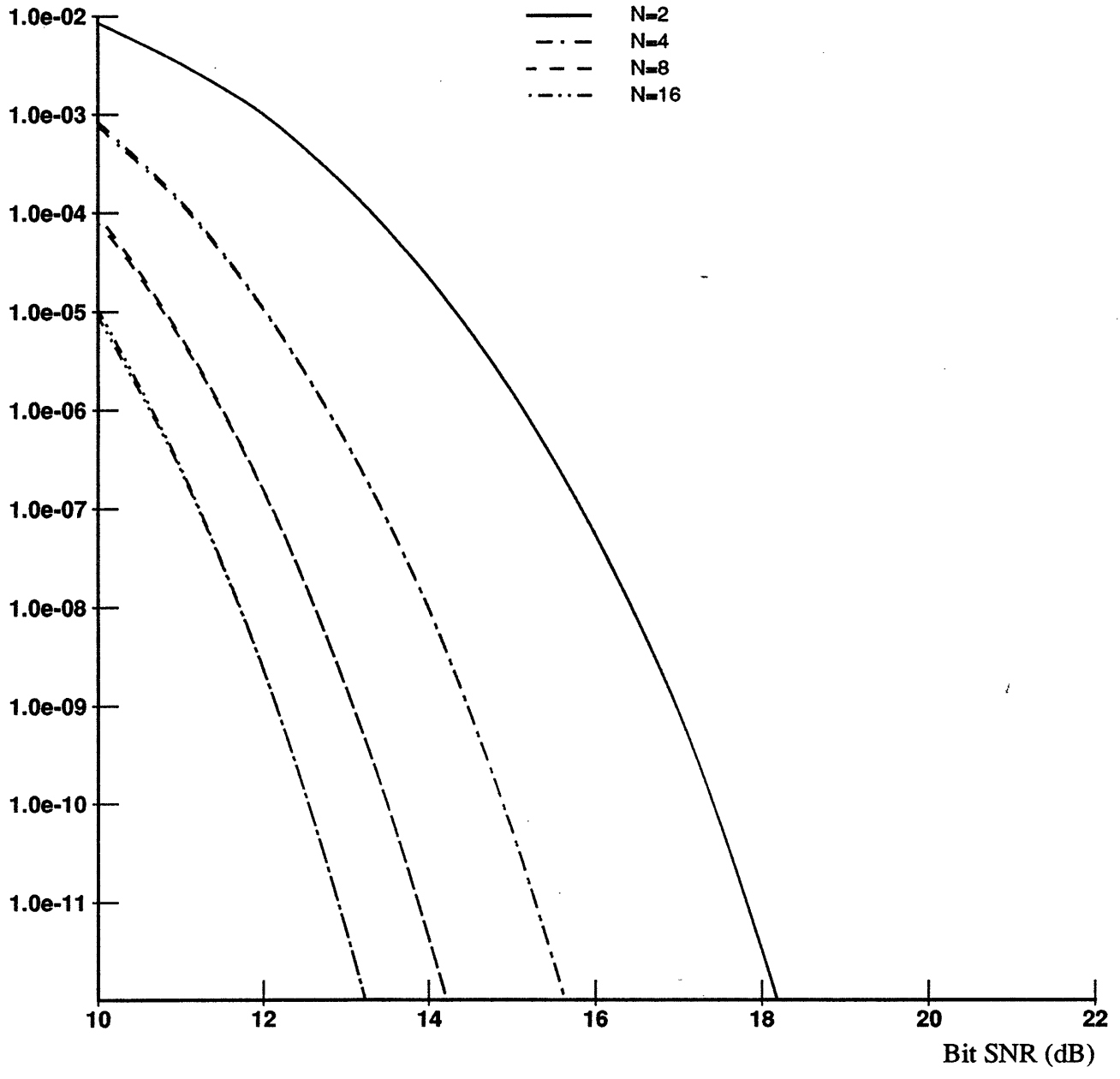

Penalty (dB)
18.0
15.0
12.0
9.0
6.0
3.0
0.0
Single sample
Multisample single-filter
Multisample double-filter
0.0
0.7
1.4
2.1
2.8
3.5
4.2
4.9
5.6
γ

Fig. 13


1.0e-02
1.0e-03
1.0e-04
1.0e-05
1.0e-06
1.0e-07
1.0e-08
1.0e-09
1.0e-10
1.0e-11
N=2
N=4
N=8
N=16
10
12
14
16
18
20
22
Bit SNR (dB)